FORMAL EVOLUTIONARY MODELING AND THE PROBLEMS OF POLITICAL SCIENCE

by

OLEG SMIRNOV

A DISSERTATION

Presented to the Department of Political Science
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

August 2005

ProQuest Information and Learning Company
300 North Zeeb Road
P.O. Box 1346
Ann Arbor, MI 48106-1346

“Formal Evolutionary Modeling and the Problems of Political Science,” a dissertation prepared by Oleg Smirnov in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Political Science. This dissertation has been approved and accepted by:

Dr. John M. Orbell, Chair of the Examining Committee

July 8 2005

Date

Committee in Charge:

Dr. John Orbell, Chair
Dr. Van Kolpin
Dr. William Mitchell
Dr. Mikhail Myagkov

Accepted by:

Dean of the Graduate School

An Abstract of the Dissertation of

Oleg Smirnov for the degree of Doctor of Philosophy

in the Department of Political Science to be taken August 2005

Title: FORMAL EVOLUTIONARY MODELING AND THE PROBLEMS OF POLITICAL SCIENCE

Approved: ________________________________
John M. Orbell

Formal theory in political science is an approach to studying political phenomena through the language of logic and mathematics. Is a *science* of politics possible? Formal political theory has lead to many advances in the discipline; however, it is facing an increasing number of challenges. I argue that some of the most serious challenges to formal theory can be addressed with a synthesis of game theory, computational modeling, and the theory of evolution. Ultimately, I claim that the science of politics is, in fact, possible if political theory is based upon a rigorous foundation of mathematics *and* if it is also systematically connected to the life sciences.

I start by examining common challenges to formal political theory, including the assumption of hyper-rationality and a related problem of incorrect predictions about human behavior. This examination suggests that the fundamental problem that formal theorists face is not methodological but rather substantive – an imperfect model of a man.

To emphasize the importance of a good underlying such model, I present the case of prospect theory – a formal scientific analysis of human behavior devoid of sound evolutionary basis (in social sciences).

To discover a good model of a man, formal theorists may have to turn to the theory of evolution. However, the solution will be more substantive than methodological—as is the problem. Evolutionary theory *is* a formal theory and it is a natural step forward for formal theorists since the logic of evolution can be expressed mathematically. I show that formal evolutionary modeling – evolutionary game theory, models of adaptive learning, and evolutionary computation (computer simulation) – can be useful for addressing some of the most challenging problems of political science without abandoning the rigor of logic and mathematics.

I apply formal evolutionary models to three different issues: endogenous agenda-setting, cooperation and altruistic punishment, heroism and intergroup violence. The models I develop are designed to provide realistic empirically testable predictions consistent with the view of human behavior now emerging from the life sciences.

CURRICULUM VITAE

NAME OF AUTHOR: Oleg Smirnov

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon
Rostov State University

DEGREES AWARDED:

Doctor of Philosophy in Political Science, 2005, University of Oregon
Master of Arts in Economics, 2005, University of Oregon
Master of Science in Political Science, 2003, University of Oregon
Bachelor of Arts in Journalism, 2000, Rostov State University

AREAS OF SPECIAL INTEREST:

Formal political theory

GRANTS, AWARDS AND HONORS:

Stephen L. Wasby Political Science Dissertation Research Award, University of Oregon, 2004-2005.

Political Economy Research Center Summer Fellowship, Bozeman, MT, 2003.

Empirical Implications of Theoretical Model Summer Training Institute, NSF Award # 0215621, Harvard University, 2002.

Empirical Implications of Theoretical Models best paper award, Midwestern Political Science Association, 2002.

Complex Systems Summer School Grant, Santa Fe Institute, 2001.

PUBLICATIONS:

Smirnov, Oleg, and James Fowler. 2005. "Dynamic Parties and Social Turnout: An Agent-Based Model." *American Journal of Sociology* 110:4, January.

Smirnov, Oleg, James Fowler, and Timothy Johnson. 2005. "Egalitarian Motive and Altruistic Punishment." *Nature* 433, 06 January, doi:10.1038/nature03256.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Table Page

CHAPTER ONE

INTRODUCTION: ROAD MAP

Evolutionary theorizing in the social sciences has been largely driven by economists with notable examples being Veblen (1898), Alchian (1950), Nelson and Winter (1973, 1982), and Hodgson (1993). Alfred Marshall and John Nash were also among the scientists who suggested possible links between evolutionary process and economic behaviour. Evolutionary game theory has a more recent history exemplified by the works of biologist Smith (1982) and economists Weibull (1995), Samuelson (1997), Hofbauer and Sigmund (1998), and Gintis (2000), which are the best systematic presentation of the theory. In political science, representation of formal evolutionary modeling is confined to a very small number of recent works such as Axelrod (1980, 1986), Kollman, Miller, and Page (1992), Bendor and Swistak (1997), and Orbell et al (2004). Broad application of formal evolutionary modeling in political science has yet to be suggested.

The goal of my research is to show that formal evolutionary modeling can provide a number of insights into the problems of political science. The tools of formal evolutionary theory can be used to model dynamic processes in which individuals have limited information-processing capacities. These tools typically relax the assumption of individual hyper-rationality and, as a result, often lead to better predictions about social behaviour without giving up the rigor of mathematics and formal theory.

Formal evolutionary theory has yet to be accepted in political science as a new, perhaps complimentary, methodology. Political science traditionally benefits from theories and methodologies developed across disciplines. Notable examples are behaviorism adopted from psychology and rational choice theory adopted from economics. Recent advances in the life sciences, however, remain largely unnoticed. One of the objective reasons for the huge gap between political science and life sciences is in the lack of methodological background and technical difficulty that political scientists inevitably face when trying to apply the tools developed in evolutionary biology, evolutionary mathematics, and evolutionary psychology. By developing these tools and applying them in the context of the discipline, we are more likely to have larger support from contemporary political scientists and to have more graduate students capable of doing the implied interdisciplinary research—thus to be more likely to bridge the gap between social and life sciences.

There *is* a formal theory in political science but it is not evolutionary in any way. Formal theory in political science is an approach to studying political phenomena through the language of logic and mathematics. In essence, formal political theory is a relatively narrow application of the tools borrowed from economics in order to study politics. Formal theory has lead to many advances in the discipline; however, it is also facing an increasing number of challenges.

I begin my project by examining some of the most serious challenges to formal theory. Among those challenges one stands out—the assumption of hyper-rationality resulting in incorrect predictions about human behavior; I will call this the "main challenge." An imperfect model of mind underlying contemporary formal theory appears to be a fundamental problem that has to be addressed if formal theorists want to carry the banner of science in the discipline. In my work, I suggest how we can approach the main challenge with a synthesis of game theory, computational modeling, and the theory of evolution.

The importance of a good underlying model of a man is emphasized in chapter 3 devoted to prospect theory (Kahneman, Slovic, and Tversky 1982). Prospect theory is now more than 30 years old and it is well-recognized in the scientific community as exemplified by the Noble Prize to Daniel Kahneman in Economics in 2002. The theory, however, is a perfect example how good scientific methodology may fail without a good substantive basis. For years the students of prospect theory believed that systematic deviations from the predictions of economic models of behavior are a result of individual irrationality, "heuristics and biases." That was *not* a good model of a man. In fact, we can even hardly call that a model. Meanwhile scholars in the life sciences observed prospect theory behavior in animals and explained *why* we observe such behavior (see Bateson 2002 for an extensive overview). Apparently risk-sensitive foraging preferences in animals are a product of an evolutionary process. Risk-aversion in gains (when survival is secured) and risk-seekingness in losses (when survival is not secured) is a behavioral trait which is selected in a fitness competition with other traits, such as the expected utility trait. In the social sciences, nevertheless, many continue to explain prospect theory behavior arguing that "people just make mistakes."

In my work, I suggest that formal evolutionary modeling may help us explain a variety of political phenomena and not repeat some of the mistakes of the prospect theory. To discover a good model of a man, formal theorists may have to turn to the theory of evolution. In fact, evolutionary theory *is* a formal theory and it is a natural step forward for formal theorists since the logic of evolution can be expressed mathematically. I consider four types of [overlapping] formal evolutionary models:

a) evolutionary game theoretic models (replicator dynamics and evolutionary stable strategies);
b) models of adaptive learning;
c) models of evolutionary imitation;
d) evolutionary simulation (genetic algorithm and related evolutionary computation).

I discuss each of the models in detail in respective chapters below but briefly they can be summarized as follows. *Evolutionary game theory* can be subdivided into the models of evolutionary stability and replicator dynamics. Evolutionary stable strategies (ESS) are behavioral types which are robust with respect to invasion of other behavioral types – in fact, ESS is nothing more than a refinement of the Nash equilibrium solution. Replicator dynamics describes evolutionary stable states and how the composition of types in the population changes *on the way* toward those states.

Models of *adaptive learning* describe cultural evolution. The gist of the approach is that individuals choose the best course of action *given* the past behavior of other individuals in the population. As the new information becomes available, individuals adjust their best-response-to-the-past behavior accordingly. In this respect, adaptive learning can be regarded as a form of Bayesian learning and it is only a step away from standard game theoretic behavior. The only difference is the absence of rational expectations about others. Instead of rational (i.e., prospective) expectations, individuals rely on past history, i.e., exercise retrospective rationality.

Models of *evolutionary imitation* are sometimes confused with models of adaptive learning. The types of models, however, are very different. In evolutionary imitation, individuals also use past history but they do so in order to myopically adopt the behavior of the most successful individuals. They neither form rational expectations about other nor make a rational decision of their own. Instead they merely copy (imitate) the best decision of another person in the past.

Finally, *evolutionary computation* is a very broad category of evolutionary models on the basis of computer simulation. Such models may be based on the logic of adaptive learning or evolutionary imitation, or they may be different. Genetic algorithms in all their huge variety are an example of a unique evolutionary algorithm (mutation – selection – crossover). Ultimately, however, evolutionary computation is simple and has a unique underlying logic: some types do better than others in a given environment, and those that do better are more likely to reproduce. The complexity arises when it comes to determining what is "doing better" – as exemplified below in the chapter on multilevel selection and the evolution of heroism.

Furthermore, there is another – more general – classification of evolutionary models. In some models the evolutionary process is an *analogy* while in others evolution is *homology*. Analogy refers to the cases when the evolutionary dynamic is an appropriate metaphor to describe the process. For example, application of evolutionary models in the context of Congressional politics is a definite example of evolutionary modeling as an analogy. On the other hand, the evolution of actual biological traits and the "wet" structure of the brain is a case of homology. In this case, formal evolutionary models *have* to approximate the environment of evolutionary adaptedness (EEA). Here I have two models which can be viewed as homology: evolution of altruistic punishment (chapter 6) and evolution of heroism as a domain-specific form of altruism (chapter 7).

Substantively, the work aspires to provide answers to some of the most difficult and interesting questions in political science: the puzzle of human cooperation and the tragedy of the commons; the logic of collective action and possibility of internal enforcement; between- and within-group competition; and a number of puzzles within the theories of electoral competition such as voters' turnout, candidates' electoral strategies, endogenous agenda-setting, and strategic voting. An attempt to give new predictions and provide alternative explanations of these phenomena should be interesting to students of politics, even those who are interested in neither formal theory nor the theory of evolution.

The questions that I ask are the standard political science questions. How can we reconcile the inescapable logic of the tragedy of the commons with apparent examples of human cooperation and altruism in general? Must the tragedy of the commons be solved exclusively either private property rights or centralized coercion? What are some other mechanisms – besides Tit-for-Tat – that may lead to cooperation in realistic settings? If individuals compete not only within a group but also with other individuals from other groups, can social cooperation emerge and be sustained? Why do people vote when the probability of being pivotal is far less than probability of death on the way to the polling place? Why do politicians offer divergent platforms when the electoral logic suggests convergence to the median voter? Although my primary focus is formal evolutionary modeling *as a methodology*, substantive results reported here should be also of some interest to the broad audience of political scientists.

A word on terminology. Throughout the presentation I refer to "formal evolutionary modeling" or "formal evolutionary theory." I intentionally emphasize this wording as opposed to "formal theory" or "evolutionary theory." Most of the formal theory in political science relies on the tools developed in economics and it does not have connections with evolutionary theory and life sciences in general. On the other hand, few evolutionary models in political science use the term “evolution” as a metaphor and typically belong to the realm of political philosophy. “Formal evolutionary theory” as I refer to it is based upon the language of mathematics and it is informed by recent advances in the life sciences, in particular, evolutionary biology and evolutionary psychology.

Advocating a substantial methodological change for political scientists and addressing the problems above is admittedly very ambitious. A major question for me is whether this endeavor is feasible and my goals realistic. I believe the answer is positive. Although advocating formal evolutionary theory as a mainstream methodology for political science is unorthodox, I am building upon a highly developed research with strong interdisciplinary roots. Critical underlying evolutionary theory and formal axioms already exist. Application of these tools within Political Science will require certain changes, but there is already a strong theoretical foundation for my research. In the end, the goal of this research is to advocate a new methodology for political science that can make realistic predictions about social processes without abandoning the rigor of formal theory and mathematics. Formal evolutionary modeling is a primary candidate for such methodology. Can evolutionary modeling be used by political scientists in their research? My task is to provide a positive answer to that question.

CHAPTER TWO

FORMAL MODELING IN POLITICAL SCIENCE AND ITS CHALLENGES

Formal theory in political science is an approach to studying political phenomena by means of mathematics and economic theory. Formal political theory has lead to many advances in the discipline, but is facing an increasing number of challenges. Here I examine some of the most common such challenges. I divide the "counterattack on the economics of politics" (Mitchell 1989) into six parts, which are:

1. The economic paradigm is irrelevant when applied in the analysis of politics.
2. This approach hampers pluralism of theories and leads to academic domination of one theory (Walt 1999).
3. Little has been learned as the parading findings are banal post hoc tautologies (Green and Shapiro 1994).
4. Universal theory of politics takes us away from understanding of the details of "richly textured life" (Fiske 1992).
5. Empirical contributions in the field are "few, far between, and considerably more modest than the combination of mystique and methodological fanfare surrounding the rational choice movement" (Green and Shapiro 1994).
6. People are not *unboundedly* rational utility-maximizers (Elster 1983; Slovic 2001; Gigerenzer and Todd 1999; Quattrone and Tversky 1988; Simon 1985; March 1978; Henrich et. al. 2001; Cosmides and Tooby 1994; Fox 1992).

I argue that the first five challenges are questionable whereas the sixth challenge is formidable and should be addressed by means of new tools and concepts. Formal evolutionary modeling and the dialog with life sciences may be the required solution.

The economic paradigm is irrelevant when applied in the analysis of politics

Gordon Tullock called the economic analysis to study democracy a "strict theory of politics" (Tullock and Buchanan, 1962) to distinguish it from traditional historico-philosophical analysis once dominating political science. Early rational choice theorists often studied the same phenomena that were the object of inquiry of the traditional political scientists—constitutional design, collective action and provision of public goods, electoral competition, democratic ideals and their implementation. The main difference was methodological. Political economists such as Arrow, Downs, Olson, Tullock, and Buchanan believed that political science should be oriented toward *explanation* and *analysis*, rather than *descriptive* and *judgmental* vision of the world. Explanation and analysis should be based on a set of general rules which can not be subjected to different interpretations. Rational choice theory emerged in order to answer the very basic but challenging question "why?" According to Shepsle, traditional ways to study politics – storytelling, thick description, political history writing, normative judgments, and the accumulation of details – did not always add up to much more than a pile of facts (Shepsle and Bonchek, 1997).

As Mitchell (1991) put it, any penetrating analysis of political interaction must account for the economic motivations of the participants. The concept of economic rationality is a cornerstone in economic theory. The term "rational" per se does not bear a normative meaning: according to Downs (1957), the term "rational" is unambiguous and never applied to an agent's end but only to her means. Rationality here implies a desire of individuals to achieve their goals in the most efficient, value maximizing way. As Buchanan put it *individualism* is both an "analytical method and a system of social order" (1962, 315). In the public choice perspective, all people who are involved in politics act out of self-interest, be they candidates or voters (Downs 1957), individuals comprising groups and organizations (Olson 1965), or individuals choosing a constitution (Tullock and Buchanan, 1962).

Tullock states that the strict theory of politics can be divided into three areas: "theory of committees and elections," (Black 1958) "theory of parties and candidates," (Downs 1957) and the "theory of constitutions." (Buchanan and Tullock 1962). However, nowadays formal political theorists study most aspects of human action within and outside political science. Furthermore, formal theory has two faces—positive (Binmore 1994; Ordeshook 1986, 1992) and normative (Mueller, 1989, 445; Buchanan, 1962, 305). "What ought to be?" is the normative question while "What is?" remains the basic positive one. As Buchanan put it, the distinction has separated the moral philosopher on the one hand from the scientist on the other. However, along with Tullock, as well as other public choice scholars, he also argues that this dichotomy is too simple. It would not be correct to label traditional political theories as "normative" and the strict theory of politics as "positive." It is true that game theory and highly abstract models of legislative bargaining or electoral competition lie in the domain of positive political theory; but there is more to public choice, however, than just providing an answer to the question "What is?" Tullock and Buchanan, for example, were accused in propagating conservative ideals. Many of public choice works were also seen as a direct challenge to the Marx's economic theory of classes. Rational choice theory has a lot of normative implications. This is where one finds it to be an offspring of both economics and traditional political science. Public choice not only posts the question "How efficient [political] markets are?" but also "How fair these markets are?" Stevens (1993, 25) outlines efficiency and equity (fairness) as primary reasons for public policy.

Formal theorists study various phenomena of political life to analyze whether they are (1) efficient and (2) fair (Becker 1976; Bartlett 1973; Breton 1974). In this respect, there is no doubt that this line of thought is highly relevant to the study of politics.

This approach hampers pluralism of theories and leads to academic domination of one theory.

There are two responses to the criticism that formal theory hampers healthy theoretical pluralism: factual and theoretical. As Martin (1999) showed, in fact, there is no rational choice theory domination as many would argue. The number of articles in academic journals (including the leading ones) is not as large as many think. The number of academic positions for which junior scholars compete is small. In sheer numbers, formal theorists can be even seen as underrepresented.

Second, there is a philosophical response: some critics miss the importance of theory in providing insights that are logically connected to one another in an integrated analytical framework. This is a necessary condition for progress in social science. Niou and Ordeshook (1999) argue that very little of what traditional researchers label "theory" is theory in any true sense, but instead is often best described as a demonstration of one's ability to "cobble" together assumptions and derive something that can be labeled "lemma" or "theorem."

Noone complains about the "lack of pluralism" in physics. Noone complains about dominance of the three laws of Newton and argues that such "discrimination" hampers scientific progress. Of course, one may note that political science, as a social science, is not close to the achievements in physics, but the point made above simply states that a domination of one theory per se does not automatically mean scientific gridlock or non-science-based discrimination. It is in the interest of all political scientists to have a coherent and practically useful discipline.

Little has been learned as the findings are "banal post hoc tautologies."

The often heard challenge that formal theorists state the obvious is a result of selected pieces of early public choice findings often presented without a context. Many works in the field are descriptive (hence, the "post hoc" criticism) but a clear understanding of causal processes is what has been missing in previous research. Moreover, such "post hoc" enterprise often generates testable hypotheses and empirical implications.

Applying the banal-post-hoc-tautologies challenge to the classical works is particularly questionable. Obviously, scholars have long been struggling to find a way to make this world and country a better place. Formally, in doing so, they were looking for a Nash-implementable social choice rule. The admirably "thick" research was rather incoherent with a lot of different interpretations and suggestions how to make democracy work under various conditions. Scholars felt that it was *difficult* to create such rule. Kenneth Arrow proved that it is not just difficult to find a rule that satisfies a few very basic democratic conditions, it is *impossible*. This is really an astonishing result. Arrow mathematically proved that there is no way to aggregate preferences without violating at least one of the five basic conditions (Arrow 1951; Inman 1987; Stevens 1993; Kelly 1988): *Pareto optimality*, *non-dictatorship*, *unrestricted domain* (the collective choice process is capable of reaching collective decision for all possible combinations of individual preference orderings of the alternatives), *rationality* (transitivity), and *independence of irrelevant alternatives* (if you prefer x to y, you choose x without taking into account z). The conditions do not, on first consideration, seem at odds with one another; they are rather basic and desirable. And as we know there are billions of social choice rules to work with. But if we cannot find a rule satisfying all of Arrow's criteria, then we certainly cannot find a social choice rule satisfying those plus some more.

I now turn to some other public choice works which are unambiguously neither post hoc nor trivial. Two seminal works in this area are *An Economic Theory of Democracy* by Anthony Downs (1957) and Duncan Black's *The Theory of Committees and Elections* (1958). The Downsian model is based on strict *economic rules* of *rational* behavior. Candidates as well as voters are assumed to be *self-interested utility-maximizing individuals*. Political interaction is straight-forward too: parties maximize their utilities by winning more votes, voters maximize their utility by choosing a better candidate. Finally, the main focus of the book is to explain why politicians behave as they do and what the consequences of such behavior are. In the words of Anthony Downs the theory was constructed in order "to provide a realistic behavior rule for a rational government and to trace its implications" (Downs 1957, 3).

The crux of the model is the representation of electoral competition by means of one-dimensional Euclidian issue space, with extreme left and right corners representing correspondingly extreme liberal and conservative candidates' platforms. The idea is intuitively appealing and its origins can be traced back to Galton (1907) and the analysis of "middlemost" position in a variety of competitive contexts; a much more elaborative model was employed later by Hotelling (1929, 41). Imagine a competition between two shops on one street. Which one would a consumer choose when everything else (service, goods, etc.) is equal? The answer is obvious – the one which is geographically closer this consumer. This would lead both shops to move their locations to the center of the street in the fight to get more customers. The electoral competition analogy is straight-forward. What later came to be known as Median Voter Theorem formally states that if all voters vote and their preferences are single-peaked on a single dimension, then the median ideal preference can defeat all other positions in a pairwise vote. As a result candidates in a two-party system converge to the median voter who is typically situated near the center of the issue space. This is another astonishing result. We may disagree with it, try to disprove, but it certainly is not banal and post hoc. Indeed, Downsian spatial model appeared to be one of the most influential ideas in political science.

Duncan Black relaxes one of the key assumptions of the Downsian model and Median Voter Theorem: what happens when voters' preferences are not single-peaked? This often leads to a violation of the transitivity of *collective* preferences, in which option A is preferred to B, B is preferred to C, but C is, in turn, preferred to A. This cycling problem is known as Condorcet paradox (who was the first to find it in 1785), or Arrow's problem. Which candidate *ought* to be elected in such case? The problem is further complicated by the fact that different voting rules sometimes lead to different outcomes, even when voters' preferences remain the same. As C.L. Dodgson—known as Lewis Carroll, the creator of *Alice in the Wonderland* fairy tales—showed, mechanisms like methods of simple majority, absolute majority, pairwise elimination, elimination of the last, method of marks (Borda count analog), and method of nomination lead to quite different electoral outcomes, which often do not represent the "actual" preferences of voters. Later

Riker (1982, 1986) confirmed that, indeed, different rules produce different results, therefore the ones who choose the rules (the power of agenda-setting) have an option of manipulating the outcomes. Furthermore, voting on the choice of the rules themselves leads to the same problem as the rules become nothing more than congealed preferences (Riker 1980). Hence, the idea of structure-induced equilibrium (Shepsle, 1979) unfortunately does not provide a complete solution to the problem of cycling majority as well.

Finally I will turn to Tullock and Buchanan in order to demonstrate one more time how non-trivial and important public choice findings are. Although the main goal of *The Calculus of Consent* was to analyze "individual choice of constitutional rules" (Mitchell 1989), the book provides an intellectual and methodological basis for almost all (or arguably *all*) research in both social choice and public choice fields.

Tullock and Buchanan provide an "efficiency explanation for the *voluntary acceptance of coercion*" (Stevens 1993, 134), why self-interested individuals might submit to binding decision rules even though some group decision may not be in their interest. Another important result is the introduction of such concepts as *decision-making costs* and *political externality costs*, later utilized by Schelling and other rational choice theory researchers. The model by Tullock and Buchanan, while intellectually powerful, is surprisingly simple for it allows one to conclude which choice of constitutions suits better various individuals and groups. "A Conceptual Classification" presented in the book is an example. Any individual is assumed to be able to order the expected costs from purely individualistic behavior, *a*; private, voluntary, but jointly organized behavior, *b*; and collective or government action, *g*. (Tullock and Buchanan 1962, 51). There are six possible orderings of the three variables. In turn, this leads to six distinct situations, in which different organization of society leads to efficient outcome and maximization of individual utility. Applying the efficiency criterion to legislative bargaining, and rent-seeking and log-rolling, the authors are able to conclude that the former is indeed a "bad thing," however, the latter is not. This was something unimaginable for traditional political theorists!

There may be a psychological explanation why traditional scholars struggle to appreciate or at least accept public choice findings. As Dennis Mueller (1989) notes, a lot of the results are either "disheartening descriptions of democratic malfunctions" or utopian proposals to improve political institutions. What he rightfully notices, though, is the fact that contemporary Switzerland or the United States also would be seen as utopia in, say, Middle Ages, or even present-day countries under totalitarian rule.

Universal theory of politics takes us away from understanding of the details of "richly textured life."

It is true that public choice scholars, like all scientists, are prone to think in *abstractions* (McKenzie and Tullock 1975), and that reducing reality to the relationships that are important and bringing the inquiry down to manageable proportions means that some of the world's complexities are inevitably lost. All models are inherently "unrealistic" in this sense, with only the real world objects and phenomena being "realistic" (although some question even that!). A map of a city is not "realistic" and some physicists argue that the model of electron is not realistic either. Nevertheless, we do buy maps and have TV and computers. Every model is a simplification of reality. As Riker (1962, 8) put it, the main advantage of a model is that it is a "convenient way of generating hypotheses and something of a brake on inconsistency." The question is whether it is adequate and useful

Traditional scholars claim that "devaluation of the contextual in favor of the universal" and exclusion of the "significance of the particular" (Fiske 1992) take rational choice theorists away from an understanding of social life, whereas in the construction of a richly textured life, social identity (or social consciousness) leads traditional scholars to comprehensive knowledge of social order and culture.

Unfortunately comprehensive knowledge of the richly textured life alone is not capable of advancing science. An accumulation of highly incoherent, and subjective facts that are vulnerable to different interpretations does not produce science since verification of judgmental hypotheses is impossible. Therefore, if a *scientist* has to choose between a model and a "richly textured life," he or she has to choose the former.

The admittedly very simple models of Arrow, Downs, Olson, Tullock, and Buchanan provide insights that wouldn't be possible otherwise, without model-building. As we already discussed, these models lead us to non-trivial and important results. Are these results empirically useful? This is the next criticism of the "counterattack on the economics of politics."

Empirical contributions in the field are "few, far between, and considerably more modest than the combination of mystique and methodological fanfare surrounding the rational choice movement" (Green and Shapiro 1994).

This a general criticism directed at the whole theory rather than at specific works. Interestingly, some of public choice's classic works do have an empirical part despite their providing a theoretical and methodological basis for further research. Many traditional political scientists, do not read (unfortunately) farther than the first three chapters of Anthony Downs and more than a non-technical summary of chapter I, section D of Olson's book. Taking a closer look at these public choice classics, traditional scholars would see that there are numerous real-politics examples. *An Economic Theory of Democracy* has several empirical accounts of spatial theory with the emergence of Labour Party in England being the most spectacular case.

The Olson's *The Logic of Collective Action* provides very important insights into the problem of provision and distribution of public goods. The dominating theories at the time assumed that individual join clubs and are engaged in collective action because of [vaguely defined] *common interests*. Olson argues that individuals often have conflicting [individual] interests, tend to free-ride and have difficulty coordinating on multiple objectives. Large groups are especially problematic for three reasons: (1) coercion is less often less effective; (2) there is a low probability of being pivotal, and (3) the factor of anonymity. The larger the group, the farther it will fall short of providing an optimal amount of a collective good.

This theoretical part is what *The Logic of Collective Action* is known for. However, this is merely a theoretical *introduction* to an empirical analysis and it is less than one-third of the book! The case studies of Labour unions and interest groups in America are usually ignored.

Nowadays only a small proportion of rational choice articles in academic journals are purely theoretical; most works include either statistical analysis or qualitative discussion of a problem. Congressional studies (Fiorina 1974; Ferejohn 1974; Shepsle 1978; Weingast 1979) and security studies (Schelling 1960; Riker 1962, chapters 7-10) present notable achievements of the economic paradigm applied in the analysis of politics, both domestic and international.

People are not unboundedly rational utility-maximizers.

This challenge, supported by experimental evidence (Kahneman and Tversky 1982; Slovic 2001; Henrich et. al. 2001), is mostly directed at the key assumptions of economists that people are maximizers of subjected expected utility, have unlimited processing power (unboundedly rational), transitive preferences, and make decision under uncertainty based on Bayesian updating. As some of the research demonstrated, all of the assumptions are fragile when people are involved in real-life (including "experimental") decision making. Paradoxically people often do not know what they want, not to mention that they often have a hard time figuring out how much they want something (utility) and how likely they are to get what they want (probability – Elster 1979, 1983).

Below I elaborate on the issue whether the absence of economic rationality means irrationality. Here I would like to make two points. First, the criticism above is more relevant to the whole field of rational choice theory in general, rather than to classic works of Public Choice by Arrow, Downs, Olson, Tullock and Buchanan. One has to start somewhere, and one usually starts with a model that, avowedly, may be a very crude oversimplification of reality. The merit of classics was that they initiated a grand research process and set up the stage for further discovery. Second, I would like to emphasize the point made by Cosmides and Tooby (1994, 327) that explicit, well-specified models of the human mind can significantly enhance the scope and specificity of economic theory. Moreover, explicit theories of the structure of the human mind can be "made endogenous to economic models in a way that preserves and expands their elegance, parsimony, and explanatory power."

Why do some of the rational choice models work and some do not? The answer may be that sometimes economists "get lucky" by creating models that are not contradictory to the structure of evolution of human brains. If this were not at least often the case, then economic models would, in general, lead to erroneous predictions.

I have considered six common challenges to formal political theory as exemplified by the works of Arrow, Downs, Olson, and Buchanan and Tullock. The first four I found groundless: Public choice works are highly relevant to the study of politics; there is no academic domination of the economic paradigm, and if there were it would not *necessary* be a scientifically "bad thing"; and finally, there is nothing wrong with scientists sacrificing some details of the "richly textured life" in order to work with more useful models. The fifth criticism (empirical implications), being important by itself, is not applicable to the classic works. Moreover, some present-day researchers are, in fact, apprehensive about the possibility of a theoretical "runaway effect" trying to implement empirical analysis as often as possible.

I consider the sixth criticism to be the most serious. Indeed, people are not as economists think about them; sometimes even the "as if" assumption does not work. To create better models, formal theorists should turn to evolutionary psychology and life sciences. As Tooby and Cosmides (1992) state, economic theory can increasingly be grounded in the theoretically and empirically derived models of human decision-making machinery that are presently being constructed within evolutionary psychology and elated disciplines. Even a methodologically sound science may be misleading if it does not have a strong evolutionary foundation. The case of prospect theory that I examine below is a perfect example of that.

CHAPTER THREE

THE CASE OF PROSPECT THEORY: WHY A GOOD THEORY NEEDS A GOOD UNDERLYING MODEL OF A MAN

"Behavioral decision theory" as exemplified by the work of Kahneman and Tversky became very popular in early 1970s with the emergence of "heuristics and biases" approach. The theory emphasizes that heuristics (mental "rules of thumb") are primarily responsible for mental errors and biased estimates that people make. Moreover, the psychologists found that deviations from laws of logic and probability have systematic character. Even though people occasionally give correct answers, the general thrust of the theory implies human irrationality and poor mental capacities. People are "cognitive misers" in a sense that they can only process little information and exercise slim cognitive capacity. Viewed through such prism, bounded rationality is equated with irrationality (Kahneman and Tversky 1972, 1973, 1979, 1982; Tversky and Kahneman 1973, 1980, 1983). This research, exemplified by the work of Kahneman and Tversky, has had a tremendous stimulating effect on the area of decision making. Moreover, the heuristics and biases (H&B) paradigm played a very important role in emergence of the competing theory of fast and frugal (F&F) heuristics (Gigerenzer 1991, 1994, 1999).

In this chapter I argue that despite its strong methodological foundation, prospect theory is fundamentally weak because it is not based on a good underlying model of man. The fast and frugal heuristics approach can be seen as a serious blow to the paradigm of Kahneman and Tversky that individuals are substantially and often irrational. This becomes clear as the argument between H&B and F&F approaches is only party factual and more interpretative; the crux of the debate is more about inferences. The "heuristics and biases approach" can be regarded as complimentary to the theory of fast and frugal heuristics, but *not vice versa.* Without an evolutionary basis, the Kahneman and Tversky paradigm is, I believe, exhausted.

Kahneman, Tversky, and other researchers distinguish several common fallacies with the following three heuristics being the most prominent: *representativeness* (judgments influenced by what is typical), *availability* (judgments based on what comes easily to mind), and *anchoring* (judgments relying on what comes first). Each of the three may bias our judgment and, thus, lead to an "incorrect" decision.

The authors define three reasons to study heuristics and biases: (1) they are "of interest of their own," (2) they can have "practical implications," and (3) the study of systematic error can "illuminate psychological processes that underlie perception and judgment." (Kahneman and Tversky 1996, p. 582). With the first reason the authors of prospect theory miss the key component of any research questions: answering the question "why?" and not only "what?" Without integrating explanation for *reasons* for certain systematic behavior, experimental data are only a collection of facts and not a theory. The second point is two-fold. One the one hand, you can improve your IQ-test score by studying and practicing problems of logic and mathematics such as Bayesian updating. On the other hand, as we discuss later, Gigerenzer and other F&F heuristics researchers show that often it pays to make "incorrect" decisions (in a normative sense)! The third point is probably the weakest part of H&B approach. After three decades of experiments and analysis, behaviorists are still "stuck with plausible yet nebulous hypotheses" (Gigerenzer, 1996, 592). Interestingly enough, Kahneman and Tversky acknowledge their dead-end themselves saying that they "were not able to offer a comprehensive treatment of the process by which different representations and different tasks evoke different heuristics." (Kahneman and Tversky, 1996, 583). Thus, all of the three reasons to study heurists and biases, as defined by the major figures of the field, fail.

Nevertheless, I will address some specific cases of alleged decision-making failure that behavioral decision theory studies. Then I will examine these from the point of view of fast and frugal heuristics and confront major conclusions of the heuristics and biases approach.

A good example of base-rate neglect is a common belief in the "hot hand" in basketball (Gilovich, Vallone, and Tversky, 1985). Psychologists mathematically show that there is no such as streak shooting even for great players as Larry Bird or Michael Jordan. The same psychological phenomenon is exemplified by the classic experimental problem "Predicting Professions." Even if it is known in advance that there is 0.75 chance that a randomly "drawn" person is engineer, people still may believe that he is a librarian on the basis of personalized information provided by the experimenter. People are not good at weighting of chance events, for example, estimating the probability that their vote will be pivotal in a national election (Quattrone and Tversky, 1988). People seem to be overconfident when they judge the probability of correctness of their answers to two-choice general knowledge questions using a one-sided rating scale. The apparent overconfidence also takes place in setting uncertainty bounds on unknown quantities (Poulton 1994). People believe that small samples are as reliable and as representative as are large samples.

Probably the most famous bias is the conjunction fallacy. In laboratory settings it is usually shown by means of the famous "Linda problem." The joint probability of A *and* B is always equal or less than probability of A, since A and B is a subset of A. For example, (a) Linda is a bank teller, (b) Linda is a bank teller who is active in the feminist movement. In experimental settings many people believe that (b) is more probable.

The grand task of the heuristics and biases approach was to understand the cognitive processes that produce such invalid judgments as well as the rare cases when the judgments are valid. The fast and frugal heuristics alternative has the *same* goal. But whereas I will argue that the latter provides an insight into the problem, the former fails miserably.

A lot of H&B and F&F heuristics experimental work overlap. In this sense (epistemologically), both teams of researchers can thought of as climbing the same mountain. Gigerenzer (1991) started the alternative conducting more or less the same experiments dealing with base-rate neglect, conjunction errors, and overconfidence. He noticed though that asking the *same* questions in terms if frequencies rather than in terms of probabilities dramatically changes results. His preliminary finding was that this method makes cognitive illusions disappear.

Gigerenzer confirmed that people use various "short-cuts" to make decisions. The profound difference was that the latter are not biases and unreliable aids, but rather useful "devices" that take advantage of the structure of information in the environment to arrive at reasonable decisions. Consider a person who has a heart attack. The doctor needs to decide quickly whether the victim should be treated as a low-risk or high-risk patient. Technically, about two dozen different medical tests are needed in order to establish a correct diagnosis. In practice, though, doctors do three simple things. First, they check if systolic blood pressure is less than 91. Second, they check if the patient is older than 62.5. Finally, they check if tachycardia is present. If *all* three questions lead to answer "yes" then the patient is high-risk; otherwise, she is not. Surprisingly this fast and frugal procedure leads to results astonishingly similar that would be achieved by means of 19 medical tests. The procedure was indeed fast: only three questions asked; and it certainly was frugal: all the choices were binary, either "yes" or "no." Thus, fast and frugal procedures serve as informational short-cuts that allow decision-maker spend less time and effort and *still* achieve comparatively good results. Fast and frugal heuristics are not a "discriminative move of nature" (using Fox's (1992) words), but beneficial tools acquired by people throughout their evolution. That is why Gigerenzer called such rationality "ecological." Our cognitive processes must have evolved to solve the kinds of problems that our ancestors faced every day. Cognitive modules are designed not for cool rationality, but for "hot cognition," to respond to "crucial events related to survival and reproduction" (Kenrick, Sadalla, and Keefe 1998, 488). Therefore, fast and frugal heuristics are genetically programmed to increase chances of the latter. Clearly, and contrary to Kahneman and Tversky, they are designed to benefit decision-maker. But people can successfully apply stereotypical thinking only in situations that resemble the ones faced by their forefathers. This is why, for example, people (as well as bumblebees!) do an analogy of Bayesian updating when they deal with frequencies in a "social" context rather than subjective probabilities in laboratory settings.

Ironically Gigerenzer made a *right* conclusion based on the *false* evidence. The irony was in the fact that frequency alone, as opposed to subjective probability, is not making people smart. I believe this is one of the reasons why Kahneman and Tversky (and others) have been so unwilling to accept a new paradigm. Behavioral theory observed systematic biases in judgments of probability *as well as* frequency (Slovic, Fischhoff, and Lichtenstein, 1982). In this regard, Kahneman and Tversky have a lot of empirical evidence that the "replacement of subjective probability judgments by estimates of relative frequency and the introduction of sequential random sampling do not provide a panacea…" (Kahneman and Tversky 1996, 585). This allowed them to call Gigerenzer's conclusion an "unqualified claim" (Kahneman and Tversky 1996, 585) and reject the whole theory. (We also should not forget about what Slovic (2001) called an "affect heuristic" – both Kahneman and Tversky *wanted* to reject the new theory as it put two decades of their research in danger).

Later Gigerenzer et al. confirmed what Kahneman and Tversky were arguing – frequencies do not matter that much. At the same with, along with other researchers (e.g., Cosmides and Tooby 1996) he found that people are smart when experimental settings in general resemble their ancestral past. After this point was made, the two sides completely left the same level of the argument. Kahneman and Tversky were still claiming that it is factual rather than interpretative. Gigerenzer on the other hand argued that the issue was not whether or not "cognitive illusions" disappear, whether or not frequency matters, but *when* and *why*. Gigerenzer's argument is much more strong since Kahneman and Tversky themselves conducted numerous experiment dealing with different framings of the same problem of decision, which lead to different mental processes. For example, they showed that graphic and verbal representations of a binomial process yield qualitatively different patterns in judgments of frequency (Tversky and Kahneman 1973), argued that the use of base-rate data is enhanced when a problem is framed as repetitive rather than unique (Kahneman and Tversky 1979), and observed that the impact of base-rate data is increased when these data are given a causal interpretation (Tversky and Kahneman 1980). Finally they demonstrated that a representation in terms of absolute frequencies largely eliminated conjunction errors (Tversky and Kahneman 1983). Therefore, the argument does not seem to be factual as Kahneman and Tversky have wanted it to be. As the argument seizes to be factual, the behaviorist paradigm no longer presents a challenge to the theory of ecological rationality, which has such fundamental basis as evolutionary theory. Heuristics and biases approach than can only serve as a small complimentary part that merely provides data for the selected cases when fast and frugal heuristics are put into an environment which is not evolutionary programmed. Clearly, Kahneman and Tversky would not have liked to see their "child-theory" to be a "working horse" collecting data for a paradigm that they did not - and then could not for some reasons - accept.

Now let's examine why and how fast and frugal heuristics "dealt a fatal blow" to the old paradigm. Consider the following classic Wason Selection Task in Table 1 (as in Gaulin and McBurney 2001; Cosmides and Tooby, 1996). The two problems are identical. Nevertheless, the first problem turns out to be much more difficult for subjects of the experiments. Evolutionary psychologists argue that the difference in difficulty is a matter of familiarity. People are evolved to solve problems framed in a social context. The two problems are identical. Nevertheless, the first problem turns out to be much more difficult for subjects of the experiments. Evolutionary psychologists argue that the difference in difficulty is a matter of familiarity.

Table 1. Wason selection task.

Problem 1			
You have been hired as a clerk. Your job is to make sure that a set of documents is marked correctly, according to the following rule: *"If the document has a D rating, then it must be marked code 3."* You have been told that there are some errors in the coding of the documents, and that you need to find the errors. Each document has a letter rating on one side and a numerical code on the other. Here are four documents. Which document(s) do you need to turn over to check for errors?			
D	**F**	**3**	**7**
Problem 2			
You have been hired as a bouncer in a bar and you must enforce the following rule: "If a person is drinking beer, then he must be over twenty years old." The cards below have information about four people in the bar. One side of each card lists a person's age and the other shows what he's drinking. Which card(s) do you need to turn to be sure no one is breaking the law?			
Beer	**coke**	**25 yrs**	**16 yrs**

The two problems are logically identical. Nevertheless, the first problem turns out to be much more difficult for subjects of the experiments. Evolutionary psychologists argue that the difference in difficulty is a matter of familiarity. People are evolved to solve problems framed in a social context. For example, cheater-detection problems are among the easiest for people even when the normative solution requires an analog of Bayesian updating. Trivers (1971) was the first to state that reciprocal altruism has been important in the evolution of our species. Interestingly, Kahneman and Tversky (1996) still found an alternative explanation for the phenomenon stating that results can be explained in terms of "extensional cues to set inclusion," but the authors prefer the "speculative interpretation that evolution has favored reasoning with frequencies but not with percentages." Again, they want to present the argument as bout frequency vs. probability, which it is not.

Gigerenzer and Hug (1992) found that only about 20 percent of college students correctly solve Wason selection tasks that involve a simple descriptive rule. At the same time about 80 percent solve the task correctly is the rule is in a social contract where someone may be cheating. The answer lies in a fundamental thesis of evolutionary psychology that human brain is *not a general-purpose processor* but rather a combination of a huge number of modules, each one designed (by evolution) to deal with specific tasks and solve specific problems. Since it was evolution that constructed this "mechanism," brain modules can only successfully deal with phenomena that was encountered in the past by our ancestors (see, for example, Kanazawa 2002). Traditional psychologists do not readily accept the idea that evolution builds specialized mechanisms. In fact, under the influence of B.F. Skinner and other behaviorists, many psychologists have believed that all learning was a single process governed by a single mechanism. On the contrary, brain is a physical system whose operation is governed solely by the laws of chemistry and physics (Tooby and Cosmides 1997). It means that all our thoughts, dreams, and feelings are produced by chemical reactions going on inside the skull. There are many cases when people had certain parts of their brains damaged (as a result of an accident). Often it causes a person to lose only a selected ability: vision, ability to walk, ability to "get inside other people's head," and sometimes even such specific abilities as face-recognition (there is a medical term for the latter that I do not recall). Therefore, the evidence that Kahneman and Tversky gathered merely fits the cases when there is no a module in our brain that would be able to solve the problem effectively.

Or maybe the normatively "correct" answer is not that correct? This is another possibility: in the real world, the gambler's fallacy is seldom a fallacy. Many people would say that a given person is likely to die within a week than within a year. Probabilistically this can't be correct. However, one can easily create a context, such as a patient already on the verge of dying, that would cause a sensible person to answer that this patient is more likely to die within a week; inferring that the question is next week versus the rest of the year, because the question makes little sense otherwise (Gigerenzer 1996). Maybe a "hot hand" of a basketball player is caused by some factors that are not included in mathematical model such as bad physical form of the defender? Or you can simply predict that the weather tomorrow will be like today, and you will be right as often as the professional meteorologists in the short run. As Gaulin and McBurney (2001, 175) put it: "If humans had evolved in casinos where their winnings translated into reproductive success, selection probably would have eliminated the gambler's fallacy. But in the real world it often pays to behave as if the past and future are not independent." In the real world it pays to be an adaptively rational Captain Kirk than an unboundedly rational First Officer Spook. If evolutionary advantage gives you an advice different from the one by logic - then "worse for the logic." (Fox 1992). Fox also makes a fascinating claim that an ultimately rational creature that would be able to explain (justify!) everything would be devoid of moral responsibility and moral blame. Consequently, it is questionable whether humans would have ever evolved (figuratively speaking, playing "defection" as the only "rational" strategy). This leads us to a re-thinking of the notion of cognitive miser, as a boundedly rational person whose fast and frugal heuristics mostly evolved in a social context (Humphrey 1976, Henrich et al. 2001, Orbell and Dawes 1991).

One can easily reject the concept of "ecological rationality" if he or she does not believe in *evolution*. In this sense, the theory of heuristics and biases is not far from the notion of original sin. It does not explain why people have such thing on the first place: after thirty years or research, the underlying processes of the three main biases are left unspecified. As Gigerenzer (1996) noted, there are two ways a theory can fail: by being wrong, or by being not even wrong, but merely indeterminate and imprecise. Heuristics and biases approach is the latter theory. Just like with the notions or original sin, Oedipus complex, etc., this theory is hard to reject for it does not have a scientific basis and at the same time potentially fits various phenomena. At the same time we can reject such theories by accepting something different, grounded on fundamental laws of science, with evolutionary dynamic being one of them. As Gigerenzer (1999, 29) put it,

> "...whereas the heuristics-and-biases program portrays heurists as a frequent hindrance to sound reasoning, rendering Homo sapiens not so sapient, we see fast and frugal heurists as enabling us to make reasonable decision and behave adaptively in our environment - Homo sapiens would be lost without them."

Prospect theory is now more than three decades old and it is well-recognized in social sciences. The theory, however, is a perfect example how good scientific methodology may fail without a good substantive basis. Risk-sensitive foraging preferences in humans (and animals) are a product of the evolutionary process (Bateson 2002). Explaining such behavior as simply irrational – "people just make mistake" – is misleading and cannot be a part of a good scientific research.

CHAPTER FOUR

POLITICAL DECISION MAKING AND EVOLUTIONARY MODELING: STRATEGIC VOTING AND ENDOGENOUS AGENDA-SETTING[1]

This chapter presents an application of formal evolutionary modeling in political science and its empirical implications for a very specific problem: voting under intransitivity of collective preferences with endogenous agendas. In this context, evolutionary methodology is an analogy, or a metaphor. Obviously, I do not argue that individuals have an innate propensity to vote strategically. Voting was not a part of the environment of evolutionary adaptedness and, therefore, humans cannot have evolved respective cognitive capacities. Nevertheless, using formal evolutionary modeling as an analogy has its own merits and provides insights which are not fully captured by the standard political science approaches. The evolutionary model describes the nature of absorbing states of each system, out-of-equilibrium dynamics, and properties of the convergence. In the end of the chapter I examine empirical implications in laboratory settings.

One of the key reasons behind the usefulness of this approach is the ability of evolutionary models to describe what happens *on the way* to equilibrium. It is implied in the standard rational choice models that individuals choose equilibrium strategies immediately. Empirical evidence suggests that this is not the case. Even when the predictions of game theory appear to be correct, it takes time and several iterations of the game for the individuals "to learn" the equilibrium (Plott 1983; Friedman 1996; Eckel and Holt 1989). The path toward equilibrium is typically a black box for game theorists. This is unfortunate since laboratory experiments and field evidence suggest that the pre-equilibrium behavior – the process - can be as interesting and important as the outcome.

---

[1] Human Subjects Compliance: Experiment protocol #E546-04, entitled "Voting on Agendas under Intransitivity of Collective Preferences," approved on April 27, 2004 by the Committee for the Protection of Human Subjects and Institutional Review Board, University of Oregon.

This is especially important when the environment is constantly changing. Moreover, in games such as Beauty Contest, following game theoretic predictions can be plain wrong and disadvantages for one's individual utility. It is almost presumed by rational (smart?) players that the game will start out-of-equilibrium. Similar dynamics are observed in public goods games and in the market experiments. Evolutionary game theory is not confined to the static refinement of Nash equilibrium known as evolutionary stable strategy (ESS). Arguably, the more interesting part of the enterprise is the analysis of adaptive dynamics and properties of convergence to the asymptotically stable states. In this paper, both models give us certain intuition about what is happening on the way to equilibrium. Although evolution is not the same as progress – a point well known in life sciences, in a static environment it can be interpreted as a search process for discovering the optimally rational strategy. Since social interaction is highly complex, not many individuals can engage in utility-maximizing behavior as prescribed by orthodox game theory. In this case, the models of adaptive learning and evolutionary change seem appealing.

Evolutionary modeling has been largely confined to the problems of collective action and the evolution of altruistic behavior. Interestingly, applications of evolutionary game theory to other political science phenomena are more likely to be found in economics where the methodology is becoming increasingly popular. Interesting examples include evolutionary game theoretic models of voting (Conley, Toossi, and Wooders 2001), fairness considerations (Guth and Pull 2004), and cultural competition (Vega-Redondo 1993). The application in this chapter suggests that evolutionary approach can be applied even in the most unlikely contexts such as asymmetric games with a small number of players and complex payoff structures.

Substantively, the paper also reports results on voting for endogenous agendas (voting mechanisms) in an experimental setting. Results show that the choice of agendas is not a mere extension of voting over alternatives, even for a set of rational and far-looking individuals, even under common knowledge of other players' preferences. Rational individuals may not have rational expectations about others, a fact with important implications for the theory of democracy, agenda-setting, and constitution-building. One of the implications is that endogenous institutional change is a complex process shaped not only by the preferences of actors over alternatives but also by the uncertainty about sophistication and idiosyncratic motives of other actors.

It has been established in the broad social science literature that different voting mechanisms, or *agendas*, often lead to different outcomes, even for the same set of players with the same preferences (Arrow 1963; Farquharson 1969; Shepsle and Weingast 1984; Banks 1985; Ordeshook and Schwartz. 1987). A well-known "evil" example illustrates how simple plurality, plurality runoff, sequential runoff, Borda count, Condorcet procedure, and approval voting – all give victory to different alternatives (Malkevitch 1990; Shepsle and Bonchek 1997). These theoretical results are supported in laboratory settings (Plott and Levine 1978; Cohen, Levine, and Plott 1978; Wilson 1986; Holt and Anderson. 1999). The experiments conducted by Plott and Levine were inspired by real life events, nicely described in the "Flying Club" story by William Riker (Riker 1986).

The fact that social preferences can be aggregated in a variety of ways – depending on the rules of the game, or agendas – is of primary importance for students of democratic theory. An obvious question is: If the agendas are so important, how are they selected in the first place? In most cases, agendas are endogenous; the rules of the game are determined in advance by the players themselves. Since (a) players have individual preferences over outcomes, and (b) agendas determine outcomes, we would expect players to favor certain agendas over others. As a result, the process of agenda-setting becomes a game in itself.

Several scholars have raised the issue of endogenous agenda-setting. Austen-Smith (1987) and Baron and Ferejohn (1987), among others, have shown that endogenous agenda formation is determinate if we assume that players are rational (in the game theoretic sense) and if they have complete information about the preferences of other players. In this case, agenda-setting is a natural extension of strategic voting over alternatives. Rational and far-looking individuals will know the outcomes of each agenda (at least probabilistically) and, in turn, this knowledge will make voting over agendas functionally equivalent to voting over alternatives. Similarly, choosing the rules for the choice of agendas pushes the problem only one step further and eventually leads to an infinite regress. Riker (1980) summarized the argument by saying that choosing a voting mechanism is not different from choosing an alternative. Agendas, from this position, become nothing more than the realization of individual goals, or "congealed preferences" as Riker put it, and not institutional mechanisms for the impartial aggregation of individual preferences. Indeed, Mouw and Mackuen (1992) provided empirical support for the view of agenda-setting as a mere extension of strategic voting by examining agenda-setting in the U.S. House of Representatives during the Eisenhower and Reagan administrations.

Rational choice models of endogenous agenda-setting are based upon the assumption of rational expectations. All voting in the models is *strategic*. Such voting behavior is fundamentally different from *sincere* voting. Ordeshook (1992) defines voting decision as "sincere" when an individual focuses only on the alternatives currently up for a vote, and chooses the one that she prefers most. To the contrary, a "strategic" voting decision is when an individual ignores the labels of the alternatives currently under consideration and chooses an alternative that yields the most preferred final consequence. Instances of strategic voting can be observed in the history of U.S. legislative politics (Enelow 1981; Riker 1986; Calvert and Fenno, Jr. 1994) and in laboratory settings (Eckel and Holt 1989). In the latter case, subjects were observed to be sincere at first but, as the game progressed, the subjects became increasingly strategic.

Along with the examples of strategic behavior there are examples of sincere voting, often in the same literature (Riker 1986). In fact, voting on the frequently cited Powell Amendment in 1956 is one of the cases when *both* strategic and sincere voting took place at the same time[2]. Some scholars are skeptical about the frequency of strategic behavior in real life (Krehbiel and Rivers 1990). Similarly, there is laboratory evidence suggesting that individuals rarely behave according to game theoretic predictions and, instead, tend to vote sincerely (Herzberg and Wilson 1988). Overall, it seems that there is no consensus as to whether voting behavior is generally strategic or generally sincere. Despite the arguments

---

[2] Some members of the Republican party strategically voted for the Powell amendment, their least preferred outcome, since they knew that the status quo, their most preferred outcome, had better chances against the amendment and not against the original bill. On the other hand, some members of the Democratic party, known as Powellians, sincerely voted for the amendment, their first choice, despite knowing that the amendment is likely to lose to the status quo, their last choice.

in favor of strategic voting, we can observe sincere voting for a variety of reasons. Even if a legislator is highly rational and far-looking, she may still vote sincerely if strategic voting is associated with an idiosyncratic cost to legislator's integrity, pressure from constituencies, or a cost to future electoral prospects. In this case, what looks like sincere voting will actually be highly rational even in the game theoretic sense, given goals other than simple expression of the legislator's policy preferences.

If voting over alternatives cannot be universally classified as strategic or sincere, the view of the agenda-setting process as a trivial extension of the voting over alternatives is questionable. If sincere and strategic voting behaviors lead to different outcomes, the selection of agendas is no longer a redundant problem. Even if a voter understands the structure of the game and votes strategically, that person may still believe that not all other voters are strategic. For example, in the well-known "Beauty Contest" experiment even the most rational individuals choose numbers that are different from the Nash equilibrium prediction (zero) due to their beliefs about other participants[3]. In fact, picking zero will rarely let you succeed in the game.

In this chapter, I examine how individuals vote for agendas. Does strategic voting for alternatives imply corresponding strategic voting over agendas? Do rational individuals who fully understand the game choose agendas which lead to their most preferred outcome when everybody else is strategic? Does the voting process correspond to game theoretic predictions? What is the role of information and communication in the game?

The results reported here suggest that the behavior of rational and far-looking individuals is imperfectly consistent with the fundamental assumption of rational expectations during the process of endogenous agenda-setting. Voting for agendas is not a redundant problem or a trivial extension of voting for alternatives. If actors are not sufficiently familiar with each other, then they may be uncertain about how strategic others are and whether there are any unforeseen factors which may influence behavior in one way or another. One actor may refuse to vote strategically in principle while another may do so

---

[3] In the Beauty Contest, the goal is to pick a number which is half the average of the numbers picked by everybody else. The unique Nash equilibrium in the game is zero.

to avoid friction with a supporting lobby. In a constant environment, an actor will eventually learn to predict the outcomes of different voting mechanisms with the process of agenda-setting indeed becoming a mere extension of voting over alternatives. However, as long as the political environment changes and congressional terms are not life-long, that should not be expected to happen.

Model

Imagine group of decision-makers – e.g., a legislative body, a board of directors, a labor union, or an academic department – divided into three parties denoted as Party A, Party G, and Party O. The parties are approximately equal in size such that no party enjoys absolute majority. The division into parties is dictated by the shared preferences of individual decision-makers. Three alternative policies are put on the table: *Alpha*, *Gamma*, and *Omega* – and only one policy can be chosen and implemented. Implementation of different policies leads to an unequal distribution of payoffs, which is common knowledge; that shown in Table 1.

Table 2. Payoff structure when the collective preferences are intransitive.

	Policy Alpha	*Policy* Gamma	*Policy* Omega
Party A	*wX*	*X*	*0*
Party G	*0*	*wX*	*X*
Party O	*X*	*0*	*wX*

Note: $X>0$, $w>1$.

The payoff structure is the classical case of intransitivity of collective preferences, known as the "paradox of voting," or "Condorcet paradox," or "Arrow's cycling problem." The group as a whole would prefer Alpha to Omega, Omega to Gamma, and Gamma to Alpha. Formally, the cycle can be expressed as $\alpha \succ \gamma \succ \omega \succ \alpha$. The game does not have a

Condorcet winner, or a policy that would beat all other policies in pairwise contests[4]. In the absence of Condorcet winner, agenda-setting becomes especially important. Three binary agendas are possible if there are three alternatives, as shown in Table 2:

Table 3. Binary voting agendas under the Condorcet paradox.

	First Round	Second Round
Agenda I	*Alpha* vs *Gamma*	Winner vs *Omega*
Agenda II	*Alpha* vs *Omega*	Winner vs *Gamma*
Agenda III	*Gamma* vs *Omega*	Winner vs *Alpha*

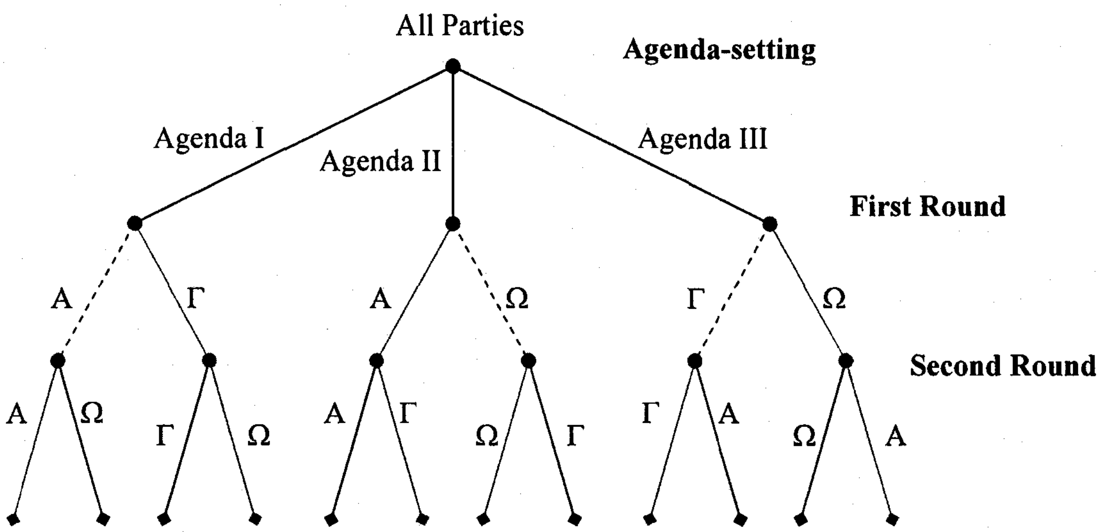
Binary agendas are typically used in legislatures. First legislators vote for the original bill versus an amendment. Then the winner is paired against a further amendment, if any, or the status quo. If several amendments are put on the table at the same time, a voting order has to be determined becoming a part of the game (see Austen-Smith 1987 for a related analysis). Since the group as whole unambiguously prefers one policy to another in pairwise contests, the order of voting becomes critical for the selection of the winner.

Classical game theory provides unambiguous predictions for the model from the assumption of rational expectations. To facilitate discussion I define two terms. A *strategic agenda* is a voting procedure that leads to your most preferred outcome if the majority of voters are strategic. A *sincere agenda* is a voting procedure that leads to your most preferred outcome if the majority of voters are sincere. The game in which parties vote for policies *and* for agendas is schematically presented in Figure 1.

---

[4] Simple plurality rule does not select a winner if the parties are of equal size: all members of party A vote for *Alpha,* members of party G vote for *Gamma,* members of party O vote for *Omega* – and we have a tie. Even if the groups are not exactly equal but an absolute majority is necessary, voting for the three policies simultaneously does not select a winner.

Figure 1. Endogenous agendas game.



Note: Each node of the game represents a collective decision by all parties. In the first node, all parties choose one of the three agendas. In subsequent nodes, all parties vote for alternatives according to the binary agenda. A stands for policy *Alpha*, G stands for policy *Gamma*, and O stands for policy *Omega*. Collective preferences are $\mathrm{A} \succ \Omega \succ \Gamma \succ \mathrm{A}$. The game has a unique subgame perfect Nash equilibrium: in the second round the parties vote sincerely (red line), in the first round the parties vote strategically (blue dashed line), in the agenda-setting round party A votes for Agenda I, party G votes for agenda III, and party O votes for agenda II.

Clearly in the second and last round of the game all voting decisions will be sincere since there are no further rounds. In the first round, sincere voters disregard the choices to be made in the second round and will vote for the labels (red line). For Agenda I, for example, sincere voters will compare *Alpha* and *Gamma* only, without consideration of their fate in the second round. Strategic voters on the other hand will ignore the labels and will be voting on the basis of the outcome in the second round (blue dashed line). For Agenda I, for example, strategic voters will compare not *Alpha* and *Gamma*, but *Alpha* and *Omega*, since *Gamma* is bound to lose to *Omega* in the second round.

According to the standard rational choice theory, individuals are far-looking. Therefore, in equilibrium all second round decision will be strategic. As a consequence, rational individuals will know that *Alpha* wins for Agenda I, *Omega* wins for Agenda II, and *Gamma* wins for Agenda III. Hence, the choice of agendas indeed becomes a trivial extension of the voting for policies. Members of each group will vote for the agenda which leads to the victory of their most preferred outcome: Party A will vote for Agenda I, party G will vote for Agenda III, and party O will vote for Agenda II.

Let us examine the actual process of finding Nash equilibrium in the game with only 3 players. First, we need to determine which alternatives win for each of the three agendas for all possible combinations of the v-V allele of the three players (Table 4).

Table 4. Possible election winners.

Players: 1, 2, 3 (v - sincere, V - strategic)	**Agenda 1 winner**: Payoffs	**Agenda 2 winner**: Payoffs	**Agenda 3 winner**: Payoffs
v, v, v	Ω : 0, X, wX	Α : wX, 0, X	Γ : X, wX, 0
V, v, v	Γ : X, wX, 0	Α : wX, 0, X	Γ : X, wX, 0
v, V, v	Ω : 0, X, wX	Ω : 0, X, wX	Γ : X, wX, 0
v, v, V	Ω : 0, X, wX	Α : wX, 0, X	Α : wX, 0, X
V, V, v	Γ : X, wX, 0	Ω : 0, X, wX	Γ : X, wX, 0
V, v, V	Γ : X, wX, 0	Α : wX, 0, X	Α : wX, 0, X
v, V, V	Ω : 0, X, wX	Ω : 0, X, wX	Α : wX, 0, X
V, V, V	Γ : X, wX, 0	Ω : 0, X, wX	Α : wX, 0, X

Second task is to establish which agenda is chosen for all possible combinations of the a-A allele of the three players (Table 5).

Table 5. The choice of agendas for each of the players.

Players: 1, 2, 3 (a - sincere, A - strategic)	Player 1 choice of agenda	Player 2 choice of agenda	Player 3 choice of agenda	Result
a, a, a	2	3	1	Tie
A, a, a	3	3	1	3
a, A, a	2	1	1	1
a, a, A	2	3	2	2
A, A, a	3	1	1	1
A, a, A	3	3	2	3
a, A, A	2	1	2	2
A, A, A	3	1	2	Tie

Now we know which agendas are chosen for a given set of a-A alleles of the three players. We also know which alternative wins for this agenda given the v-V alleles of the three players. This information allows us to construct a normal form of the game. In Table 6, player 1 is a row player, player 2 is a column player, and player 3 chooses across the four tables.

The game has a unique Nash equilibrium: (VA; VA; VA). To arrive to this result the players would have to conduct formidable analytical task captured by the previous analysis. First, the players have to represent the situation in form of a model, which we have done through the Tables 4, 5, and 6. Second, the players have to examine all 64 outcomes in order to establish which strategies are the best response strategies. Equilibrium is a combination of strategies such that no player has an incentive to deviate unilaterally, or a combination of best responses. Even with only three players, it is questionable if individuals are capable of carrying out such information-processing and analytical tasks. When the number of players increases, say, to the order of thousands, no individual will be able to deal with the resulting combinatorial explosion. A game theoretic model becomes implausible.

Table 6. Normal form of the voting game with endogenous agendas.

Pl. 3:va	**va**	**vA**	**Va**	**VA**
va	$\frac{wX+X}{3},\frac{wX+X}{3},\frac{wX+X}{3}$	$0, X, wX^*$	$\frac{X}{3},\frac{wX+2X}{3}^*,\frac{2wX}{3}$	$0, X, wX^*$
vA	$X, wX^*, 0$	$0, X, wX^*$	$X^*, wX^*, 0$	$0, X, wX^*$
Va	$\frac{wX+2X}{3}^*,\frac{2wX}{3},\frac{X}{3}$	$X^*, wX^*, 0$	$\frac{2X}{3},\frac{2wX+X}{3},\frac{wX}{3}$	$X^*, wX^*, 0$
VA	$X, wX^*, 0$	$X^*, wX^*, 0$	$X^*, wX^*, 0$	$X^*, wX^*, 0$

Pl. 3:vA	**va**	**vA**	**Va**	**VA**
va	$wX^*, 0, X$	$wX^*, 0, X$	$0, X^*, wX^*$	$0, X^*, wX^*$
vA	$wX^*, 0, X$	$\frac{wX+X}{3},\frac{wX+X}{3},\frac{wX+X}{3}$	$X^*, wX^*, 0$	$\frac{X}{3},\frac{wX+2X}{3},\frac{2wX}{3}$
Va	$wX^*, 0, X^*$	$wX^*, 0, X^*$	$0, X^*, wX^*$	$0, X^*, wX^*$
VA	$X, wX^*, 0$	$\frac{wX+2X}{3},\frac{2wX}{3},\frac{X}{3}$	$X^*, wX^*, 0$	$\frac{2X}{3}^*,\frac{2wX+X}{3},\frac{wX}{3}$

Pl. 3:Va	**va**	**vA**	**Va**	**VA**
va	$\frac{2wX}{3},\frac{X}{3},\frac{wX+2X}{3}^*$	$0, X^*, wX^*$	$\frac{wX}{3},\frac{2X}{3},\frac{2wX+X}{3}$	$0, X^*, wX^*$
vA	$wX^*, 0, X^*$	$0, X^*, wX^*$	$wX^*, 0, X^*$	$0, X^*, wX^*$
Va	$\frac{2wX+X}{3},\frac{wX}{3},\frac{2X}{3}$	$X^*, wX^*, 0$	$\frac{wX+X}{3},\frac{wX+X}{3},\frac{wX+X}{3}$	$X^*, wX^*, 0$
VA	$wX^*, 0, X^*$	$X^*, wX^*, 0$	$wX^*, 0, X^*$	$X^*, wX^*, 0$

Pl. 3:VA	**va**	**vA**	**Va**	**VA**
va	$wX^*, 0, X$	$wX^*, 0, X$	$0, X^*, wX^*$	$0, X^*, wX^*$
vA	$wX^*, 0, X^*$	$\frac{2wX}{3},\frac{X}{3},\frac{wX+2X}{3}$	$wX^*, 0, X^*$	$\frac{wX}{3},\frac{2X}{3}^*,\frac{2wX+X}{3}$
Va	$wX^*, 0, X^*$	$wX^*, 0, X^*$	$0, X^*, wX^*$	$0, X^*, wX^*$
VA	$wX^*, 0, X^*$	$\frac{2wX+X}{3},\frac{wX}{3},\frac{2X}{3}^*$	$wX^*, 0, X^*$	$\frac{wX+X}{3}^*,\frac{wX+X}{3}^*,\frac{wX+X}{3}^*$

Note: Best responses are denoted by stars. The game has a unique Nash equilibrium (VA; VA; VA).

Is it possible for a rational individual to vote for an agenda different than the one prescribed by game theory? The answer is yes, if she believes that others may vote sincerely for policies.

As an alternative to the game theoretic approach above, one can adopt a decision-theoretic framework, as follows. Assume that p_i is a subjective belief of player *i* about the probability that the majority of voters are strategic. As a result, each agenda promises a subjective expected utility depending on the beliefs of the individual. It is not difficult to establish that in the model each individual chooses among the agendas which promise the following payoffs:

Table 7. Agendas' outcomes under strategic and sincere voting.

	Strategic voting (p)	Sincere voting ($1-p$)	Expected utility
Strategic Agenda	wX	X	$wXp + b(1-p)$
Sincere Agenda	0	wX	$wX(1-p)$
Dominated Agenda	X	0	Xp

Note: $X>0$, $w>1$. For members of different parties different agendas are strategic, sincere, and dominated. For example, for a member of party A, strategic agenda is Agenda I, sincere agenda is Agenda III, and dominated agenda is Agenda II. For a member of party G, strategic agenda is Agenda III, sincere agenda is Agenda II, and dominated agenda is Agenda I.

The dominated agenda promises a payoff which is lower than the strategic agenda's payoff regardless of how others vote for policies -- sincerely or strategically. On the other hand, the choice between the strategic and sincere agendas depends on the value of p. It is trivial to show that an individual *i* prefers the strategic agenda if $p_i > (w-2)/(2w-1)$, and the sincere agenda otherwise. Thus, decision-theoretic framework allows for a rational individual to vote for a sincere agenda depending on the beliefs of the individual and the parameters of the game. Is there a basis to believe that others may vote sincerely for policies? Sometimes, there is. For example, members of the Congress may vote sincerely if the issue receives a lot of publicity and there is a cost to integrity associated with strategic voting. This and other factors often absent in our models may cause a large number of rational individual to vote sincerely. In this case, p_i is rightfully high and the choice of the sincere agenda can be justified.

Adaptive Learning

Adaptive learning relaxes some of the assumptions of the standard rational expectations model. According to the adaptive learning model, an individual follows the history of how other players have played in the past, and chooses a strategy for the future that is a best response to the past play of others (Gintis 2000). This is a pseudo-rational alternative to the approach above. Now individual task becomes easier: instead of forming rational expectations for all 64 contingencies of the game and finding a Nash equilibrium, a player only has to observe the past choices of the other two players. As a result, she only compares 4 alternatives: whether she is better off being vv, Va, vA, or VA, given the observed behavior of others.

The dynamic process of moving from one state of the game to another is now a Markov chain. For example, if all players choose strategic agenda but only one of the players votes strategically the game is in the state S0 in Figure 2. The state S0 can only go to states S1, S2, and S3 with positive probabilities which follows from the payoff structure in Table 5. Complete picture of Markov chains for the game is presented in Figure 3 (the probability transition matrix is omitted). It is obvious that (VA; VA; VA) is the unique absorbing state: once it is entered, the players stay there forever. The game has three cycles (bold blue arrows in Figure 3): with positive probabilities players can be moving away from the equilibrium which implies that it will take time for the system to converge to the absorbing state. Complete convergence can be achieved only in the limit. The adaptive learning model provides us with the same prediction as the game theoretic model despite the fact that we relax the rational expectations assumption. In addition, it provides an insight into the pre-equilibrium dynamics for all possible states of the world. In fact, using the probability transition matrix we can establish the probability of the system being in a state X at a time period T.

Figure 2. Adaptive Learning. Example: vvV, AAA.

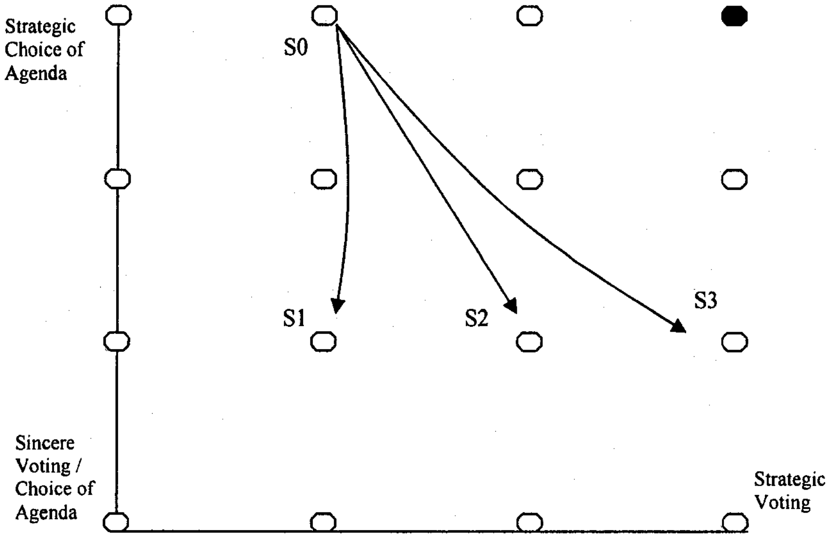
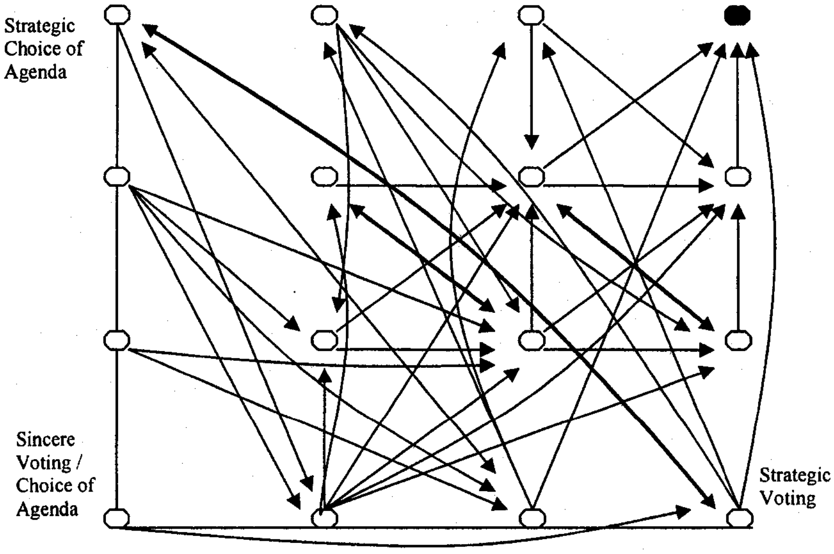


Figure 3. Adaptive Learning. Complete Picture.



Note: Bold blue arrows represent cyclic transitions which occur with positive probabilities.

Evolutionary Imitation

The players do not have to find Nash equilibrium according to the adaptive learning model, however, they have to construct a normal form of game (Table 6) – another formidable task. For complex games with many players, it would be very appealing to relax this assumption as well. This can be accomplished by means of evolutionary models such as replicator dynamics. According to these models, players blindly imitate the behavior of those who are more successful in the population. An evolutionary interpretation of replicator dynamics is that more successful types have higher fitness than others and, thus, more offspring than others.

Although there are only three players in the game, we can construct a model which is inspired by the notion of replicator dynamics and has certain parallels to it. Even with three players we can assume that the worst off individuals are likely to imitate the behavior of the better off individuals. If all players are equally successful we are in a steady state. A steady state is asymptotically stable if any [genetic/behavioral] drift does not lead to an escape from that state. The stability exists when any adjacent non-steady state behavior provides lower fitness for a player. On the other hand, if a drift leads to a high fitness behavior, the steady state is not asymptotically stable and is quickly abandoned.

These evolutionary dynamics can be represented graphically. Figure 4 represents the evolutionary dynamics of the voting game. The players are myopic and they only imitate behavior of the most successful player. The system has 4 steady states: in the four corners the players have equal utilities. Only one state is asymptotically stable: (VA; VA; VA). How can some of the game states lead to more than one other state in an evolutionary model? This is not a mistake: the reason behind that is in the specific payoffs in Table 1. For example, examine a state when there are two players who vote strategically and only one player who chooses strategic agenda (two “V”, one “v”, one “A”, two “a”).

Figure 4. Evolutionary Imitation.

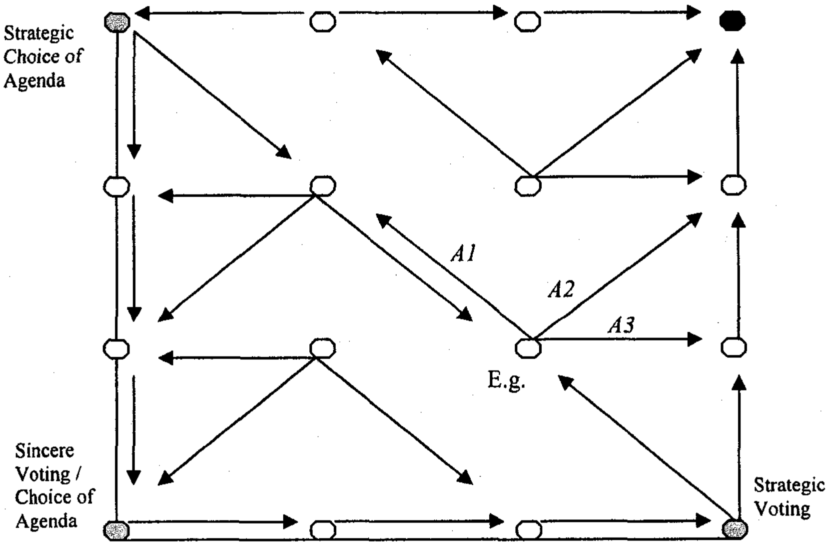


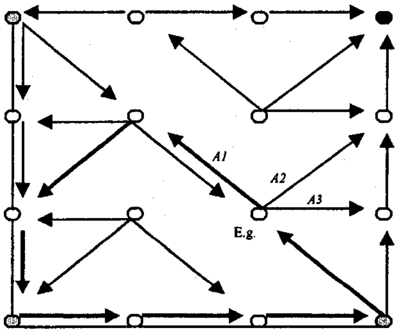
Table 8. Example: 9 possible scenarios.

	aaA	*aAa*	*Aaa*
VVv	Outcome: $0, X, wX$ Change: $Va \rightarrow vA$ *(A1)*	$X, wX, 0$ $va \rightarrow VA$ *(A2)*	$X, wX, 0$ $va \rightarrow Va$ *(A3)*
VvV	$wX, 0, X$ $va \rightarrow Va$ *(A3)*	$X, wX, 0$ $Va \rightarrow vA$ *(A1)*	$wX, 0, X$ $va \rightarrow VA$ *(A2)*
vVV	$0, X, wX$ $va \rightarrow VA$ *(A2)*	$0, X, wX$ $va \rightarrow Va$ *(A3)*	$wX, 0, X$ $Va \rightarrow vA$ *(A1)*

Note: "aaA" denotes "Player 1 chooses sincere agenda, Player 2 – sincere agenda, Player 3 – strategic agenda. "VVv" denotes "Player 1 votes strategically, Player 2 votes strategically, Player 3 votes sincerely. For a given combination of alleles of the three players, a specific outcome is realized. Given the payoffs associated with the outcome, the worst off individual changes imitates the behavior of the most successful individual. $X > 0, w > 1$. A1, A2, and A3 describe the dynamics in Figure 4. Notice that A1, A2, and A3 each take place in 3 out of 9 scenarios and, therefore, have equal probability $p = 1/3$.

The evolutionary imitation model ultimately provides us with the same prediction as the game theoretic and adaptive learning models: all players vote strategically and choose strategic agendas. At the same time, we no longer require the players to formulate the situation as a game exemplified by Table 6. Individuals adopt the same optimal behavior without having to form rational expectations about others and without solving the game for Nash equilibrium. In addition the model sheds some light on the general out-of-equilibrium dynamics. If individuals begin the game by voting strategically and by choosing a random agenda, the counter-clock dynamics in the game suggest the following. When sincere behavior prevails, the immediate improvement in one's strategies is to choose a sincere agenda: the lower left steady state is achieved. Eventually the individuals will discover that strategic behavior is more advantages and the system will transfer to the lower right steady state. In this state, however, it is clear that the choice of the sincere agenda is maladaptive. Now, the immediate improvement in one's strategy is to choose a strategic agenda. Once the system reaches the upper right corner, it stays there forever. Strategic voting and strategic agendas form an asymptotically stable equilibrium. No mistakes, random moves (mutations), deviations from the equilibrium will lead to an escape from this state. Another result, not captured by standard game theoretic models, is the possibility of a cycle (Figure 5).

Figure 5. Example of a cycle in model of evolutionary imitation.



The cycles are possible due to the by-product effect. In the lower right steady state, the most advantages mutation (change) is the change from sincere to strategic agenda. It is possible, however, that a mutant will be vA (sincere voting, strategic behavior). This type will be selected and replicated by others due to the success of the a-A allele. Interestingly, sincere voting will be imitated as well as a by-product of the selection process. Although unlikely, similar phenomenon can occur again. In this case, we go back to the mid-left part of the figure.

The three models examined here – game theoretic, adaptive leaning, and evolutionary imitation – provide us with the same general prediction: strategic voting, strategic agenda. At the same time, the models have their idiosyncrasies:

Table 9. Systematic comparison of analytical approaches.

	General Assumptions	Out-of-equilibrium dynamics	Ultimate Prediction
Game Theory (hyper-rational individual)	- Ability to formulate situation as a game in normal form - Ability to form rational expectations for other players and solve the game for Nash equilibrium	- None	- Convergence to (VA; VA; VA)
Adaptive Learning (semi/pseudo-rational individuals)	- Repeated Game - Ability to learn about the success of other players - Ability to formulate situation as a game in normal form	- General counter-clock dynamics with dramatic changes in different directions and three small cycles	- Convergence to (VA; VA; VA)
Evolutionary Imitation (myopic individuals)	- Repeated Game - Ability to learn about the success of other players	- General counter-clock dynamics with gradual changes and a possible big cycle	- Convergence to (VA; VA; VA)

To test predictions of different theoretical approaches, I conducted an experiment, in which subjects voted for policies *and* for agendas. One challenge that we face is the number of subjects. To approximate a large legislative body we need an experimental session with more than the usual 20+ subjects present in the laboratory at the *same* time. If we attempt to model the U.S. Senate, for example, we will need about a hundred subjects simultaneously present. Normally, such an endeavor is logistically impossible with standard laboratory constraints.

Another challenge associated with such an experiment is the pool of subjects: we need sophisticated participants who fully understand the game, who can spot intransitivity of collective preferences, and who are aware about the notion of strategic voting. Only subjects with such characteristics would approximate real life decision-makers, legislators, and business executives -- those who are endowed with agenda-setting authority. Normally, such pool of subjects is not available to an experimenter; it is difficult to teach subjects the paradox of voting and strategic voting during the introduction to an experiment.

Fortunately, both the logistical problem and the issue of subjects' sophistication were solved as I happened to be teaching a large upper-division class which served as an introduction to game theory in political science[5]. The experiment took place in the classroom during the seventh week -- by which time students were familiar with the basic apparatus of game theory, backward induction, the notion of strategic voting, and the paradox of voting[6]. Thus, the pool of subjects represented not a random sample of college undergraduates but rather a sample of individuals who were relatively savvy about the intricacies of voting and who (for the most part, at least) understood the importance of agenda-setting. Thus, the subjects were biased in favor of the game theoretic model but, happily, such a bias was exactly the requirement of the experiment.

---

[5] Experiment protocol #E546-04, entitled "Voting on Agendas under Intransitivity of Collective Preferences," approved on April 27, 2004 by the Committee for the Protection of Human Subjects and Institutional Review Board, University of Oregon.

[6] Given the feedback that I received after the experiment and the student performance on the subsequent tests, the experiment turned out to be not only very enjoyable but also useful for students' understanding of the subject matter.

Eighty six subjects were divided into three parties of approximately equal size: party A, party G, and party O[7]. In the experiment, all individuals were choosing among the three policies: *Alpha*, *Gamma*, and *Omega*. Realization of a policy led to the payoff structure in Table 1 with a=3, b=2, and c=1.[8] The policy was to be chosen through a binary agenda. The list of all available binary agendas is represented in Table 2. Subjects made voting decisions for *each* of the three agendas: they were told that one of the three agendas would actually be chosen by themselves through a modified plurality rule.[9] Since the parties were not exactly equal in size (31, 27, and 28 members respectively), in the interest of fairness the modified plurality rule required a minimal majority of 32 members (37%). If the required majority were not achieved, the agenda would be chosen by a random draw. Otherwise, the agenda that received the most votes would be selected as the voting mechanism.

Up to this point in the experiment, no communication was allowed. Once everyone made a voting decision with respect to agendas, the final stage of the experiment began – 15-minute party caucuses. The subjects were reseated according to their party membership and now they were allowed to discuss their voting behavior and choice of agendas. There were no constraints on the nature of the discussion as long as it took place within a party. Moreover, participation in the discussion itself was completely optional. Once the discussion was over, subjects are asked to vote for the *agenda* again. The decision remained anonymous and each individual could or could not change an original selection. Thus, the discussion merely served as an information exchange within a group of individuals with similar preferences, and did not force an individual to behave contrary to her own beliefs.

---

[7] Individuals are selected into groups alphabetically. Those whose last name begin with A-F form party A (31 members), G-N – form party G (27 members), and O-Z – form party O (28 members).

[8] The payoffs correspond to the amount of extra credit that the students receive in the class: roughly 1 point corresponds to 1 percent of the grade for the class. Personal communication with the students suggests that extra credit represent a higher value than typical experimental monetary rewards of about $20.

[9] In addition, they also have to rank their choices using a simple Borda count: 2 for their most preferred agenda, 1 for the second best, and 0 for the least preferred. This information does not have an effect on the outcome of the experiment but it provides a detailed individual preference ordering over the three agendas.

For verification purposes, each participant was asked to report her party membership on the game card (see footnote #6) and no one failed the task (see Appendix 2 for details). After that subjects were asked whether they saw a voting cycle in the table of payoffs: 87% answered positively. The final task was to express the intransitivity of collective preferences formally, $A \succ \Omega \succ \Gamma \succ A$, and 57% correctly completed the task; hereafter the subsample is referred to as "formal" whereas the other 43% form a "myopic" subsample[10].

Detecting strategic voting in the experiment was straight-forward. For each agenda one and only one of the parties faced the following dilemma: (a) vote sincerely for their first choice which was bound to lose in the second round, or (b) ignore the labels and vote strategically for their second choice which was bound to win in the second round. Party A faced the dilemma for Agenda II, for party G it was Agenda I, and for party O it was Agenda III. Those subjects who resolved the dilemma in favor of the option (a) were sincere; those who chose option (b) were strategic.

64 out of 86 subjects (74%) voted strategically for policies. Table 4 represents detailed breakdown of voting decisions for each agenda.

Table 10. Voting over policies.

	Results	Sincere Prediction	Strategic Prediction
Agenda # 1 Dilemma for Party G	A : 49 Γ : 37	A : 31 Γ : 55	A : 58 Γ : 28
Agenda # 2 Dilemma for Party A	A : 36 Ω : 50	A : 58 Ω : 28	A : 27 Ω : 59
Agenda # 3 Dilemma for Party O	Γ : 46 Ω : 40	Γ : 27 Ω : 59	Γ : 58 Ω : 28

---

[10] Note a possibility that some subjects in the myopic sample fully understood the game but simply chose not to answer this question since it had no effect on their payoffs.

Those participants who could formally express the intransitivity of collective preferences, the "formal" sample, were more likely to vote strategically for policies (bivariate probit coef. = 0.671, z = 2.24). Similarly, using the two-sample Wilcoxon rank-sum (Mann-Whitney) test we can reject the null hypothesis that formal and myopic subsamples voted for policies in the same manner (p = 0.024)[11].

When voting over agendas, 48 out of 86 subjects (56%) chose strategic agendas, 36 subjects (42%) chose sincere agendas, and only 2 subjects were clearly confused and voted for dominated agendas. Strategic voters were more likely to pick strategic agendas than sincere voters (binary probit coef. = 0.667, z = 2.11; two-sample Wilcoxon rank-sum (Mann-Whitney) test: p = 0.034). Nevertheless, 36.5% of the strategic voters *did* vote for the sincere agendas (see descriptive statistics in Appendix 1 for details). A possible explanation of this seeming paradox may be found in the decision-theoretic model above. Perhaps, rational and far-looking individuals voted for sincere agendas because they did not believe that the majority of *others* were rational and far-looking. As previously established, an individual would be better off choosing sincere agenda if p_i < (a-b)/(2a-b-c). In the experiment, a=3, b=2, and c=1. Therefore, strategic voters picking the sincere agenda must have believed that $p < 1/3$, or that there is less than 33.3% chance that the majority of voters are strategic.

An obvious explanation for the choice of agendas may be misunderstanding of the game. Perhaps subjects who could formally express the intransitivity of collective preferences voted for the strategic agendas whereas those who could not do that voted for the sincere agendas. But this was not the case. *Seeing intransitivity does not lead one choose a strategic agenda* (Table 5). The coefficient is not significant and its sign is actually negative. Alternatively, using the two-sample Wilcoxon rank-sum (Mann-Whitney) test we *fail* to reject the null hypothesis that formal and myopic subsamples voted for agendas in the same manner (p = 0.557.

---

[11] H_0: strategic voting for policies ("myopic" subsample) = strategic voting for policies ("formal subsample"). H_0 can be rejected, z = -2.25, Prob > |z| = 0.0244.

Table 11. Dichotomous dependent variable: voting for strategic agenda.

	Coef.	*Std. Err.*	*z*	*P>z*
Ability to express cycle formally (1,0)	-0.347	0.293	-1.180	0.237
Strategic voting for policies (1,0)	0.778	0.334	2.320	0.020
Constant	-0.234	0.292	-0.800	0.424

Note: Excluding the "strategic voting for policies" variable does not affect the sign and significance of the other explanatory variable.

The result indicates that the choice of sincere agendas cannot be attributed to confusion about the game. The fact that 36.5% of *strategic* voters chose *sincere* agendas further supports the point that the choice of such agendas was a *deliberate and rational attempt to maximize individual payoff given individual beliefs about the behavior of others.*

15-minute party caucuses had an expected effect on the final choice of the agendas. The number of subjects who voted for strategic agendas increased from 48 to 73 (85% of the total sample). Recall that according to the decision-theoretic approach the choice of strategic agenda is warranted if there is at least 34% chance that the majority of voters are strategic. In reality, 74% voted strategically for policies. The party caucuses gave subjects a better estimate of p, and indicated that the beliefs about low p could not possibly be supported. As a result, 25 subjects changed their mind and picked strategic agenda. Interestingly, even some of the sincere voters were persuaded by the members of their party to pick strategic agenda. The 2 subjects who initially voted for the dominated agendas changed their mind after the group discussion as well. Table 6 represented detailed breakdown of the vote for each party before and after the caucuses.

Table 12. Voting for agendas before and after party caucuses.

		Initial Choice of Agenda	Choice After Discussion	Sincere Prediction	Strategic Prediction
Party A (n = 31)	Agenda I	19	**24**	0	31
	Agenda II	0	0	0	0
	Agenda III	12	**7**	31	0
Party G (n = 27)	Agenda I	1	0	0	0
	Agenda II	12	**6**	27	0
	Agenda III	14	**21**	0	27
Party O (n = 28)	Agenda I	10	0	28	0
	Agenda II	17	**28**	0	28
	Agenda III	1	0	0	0

Notice that party O managed to deliver a unanimous vote for the strategic agenda. The other two parties moved in the same direction but failed to achieve complete unanimity.[12]

In this chapter, I examined a very specific political science problem: voting for endogenous agendas in an experimental setting. Game theory predicts that all individuals will vote strategically for policies and will also choose agendas that lead to their preferred outcome *given the strategic behavior of others*. A decision-theoretic approach, however, predicts that individuals will vote for the strategic agendas only if they believe that others are strategic. Those who do not believe in the strategic behavior of others will rationally vote for the sincere agendas. In fact, 35.6% of strategic voters, nevertheless, chose sincere agendas. This choice of agendas cannot be attributed to not understanding the game; in fact, those subjects who could formally express the voting cycling were *less* likely to pick strategic agendas.

These results have important implications for the theory of democracy and agenda-setting. Even under complete information about the preference of others, it is not always possible to know the outcome of a given agenda since individual behavior may be sincere and not strategic as our models predict. This may happen for a variety of reasons. In the experiment, it was primarily skepticism about the sophistication of others and ability to

---

[12] In case you wonder, Agenda II received the necessary majority of votes and the policy *Omega* won. Thus, the most unified party achieved their most preferred outcome.

deal with the complexity of the game. In legislatures, sincere behavior may be caused by pressure from constituencies, while in businesses it may be a part of reputation-building.

Regardless of the reason, the possibility of sincere behavior leads to an ambiguity with respect to outcomes of different voting mechanisms. Common knowledge of the players' preferences over alternatives is, therefore, a necessary but not sufficient condition for making predictions about the outcomes of different agendas. A number of external, often idiosyncratic, factors affect voting behavior of individuals although incorporating those factors into a formal model is clearly a challenge.

In light of these findings it is perhaps not surprising that William Riker referred to heresthetics ("structuring the world so that you can win") as an *art* (Riker 1986). However, as Riker certainly would have insisted, this does not mean that we should give up constructing a science of politics. But political *scientists* will make a step forward if they acknowledge that voting for agendas is more than just an extension of voting for alternatives and that it requires further modeling efforts and extra care.

In addition to the substantive results, the chapter also suggests that individual behavior is better approximated by the evolutionary models of adaptive learning. Subjects appear to make rational decisions but they do *not* form rational expectations about others. Ultimately, these models lead to the Nash equilibrium outcome, however, the behavior *on the way* to this outcome may be different from the game theoretic predictions.

CHAPTER FIVE

APPLYING EVOLUTIONARY MODELING IN THE CONTEXT OF ELECTORAL COMPETITION AND VOTERS' TURNOUT

In this chapter, I show that the models of adaptive learning and evolutionary imitation may help us tackle perhaps the most challenging problems in the political science: explaining voters' turnout and candidates' divergence. I develop an agent-based model of dynamic parties with social turnout built upon the two types of formal evolutionary models. Politicians behave according to the model of adaptive learning: they make rational decisions given the past history. Voters behave according to the model of evolutionary imitation: they copy the most successful behavior of other voters.

Substantively, this model yields significant turnout, divergent platforms, and numerous results consistent with the rational calculus of voting model and the empirical literature on social turnout. In a simplified version of the model I show how a local imitation structure inherently yields dynamics that encourage positive turnout. The model also generates new hypotheses about the importance of social networks and citizen-party interactions.

For the past half century social scientists - political scientists, economists, sociologists, and psychologists - have been intrigued by two important empirical regularities: why people vote and why political parties behave the way they do. The fact that millions of people vote may not seem to be puzzling (Berelson et al. 1954; Mackie & Rose 1997). However, given standard assumptions about rationality voter turnout cannot be easily explained. Numerous formal attempts to explain turnout predict vanishingly small turnout since the probability of affecting the outcome of an election approaches zero in large populations (Palfrey and Rosenthal 1985; Aldrich 1993; Myerson 1998). In fact, people going to polls have a much higher chance of getting into a car accident. This has led many scholars to infer that rational explanations of turnout must rely on an additional benefit derived from fulfilling a sense of duty or a general taste for voting (Riker and Ordeshook 1968).

The second interesting empirical regularity is the way political parties choose their electoral platforms – candidates offer voters policies that diverge significantly from the median voter and remain relatively stable over time (Peltzman 1984; Grofman et al 1990; Poole and Rosenthal 1984). Again, given standard assumptions about rationality this is not obvious since early models of party platforms predicted convergence to the median voter (Downs 1957; Davis et al 1970) or divergence across the entire policy space with any platform possible (McKelvey 1976; Plott 1967; Schofield 1983).[13] Subsequently, scholars explored the impact of uncertainty on policy-motivated parties (Wittman 1977; Calvert 1985; Roemer 2001). These models do yield equilibria with divergent policies, but analysis in a closed-form is rather complex. It quickly becomes intractable under all but the most basic assumptions (Roemer 2001).

Because of the complexity involved in modeling both parties and voters, past efforts have not combined them (Osborne 1995). Models of voter turnout have usually relied on assumptions of fixed party platforms, while models of platform choice have assumed a fixed level of voter turnout (usually 100%). The interdependence between people and politicians also has a dynamic character that is missing from many models because they consider a single election in isolation. Most elections are, in fact, part of a longer process of party competition and take place in a context of information about previous elections.

Economists and political scientists have also frequently abstracted away from elements that sociologists and psychologists believe to be critical for determining electoral behavior. For example, many models of elections have avoided situating voters in social networks, or social context in general. Voters are often assumed to exist independently of one another in spite of a growing body of sociological evidence suggesting that how they are situated in relation to one another plays a critical role in the decision to vote (Lazarsfeld

---

[13] The idea of platform convergence on a single issue space was borrowed from economics. Two shops fighting for customers on a single street will choose their locations in the middle of the street in order to minimize the average distance to the shop for all potential customers (Hotelling 1929).

et al 1948; Berelson et al 1954; Campbell et al 1954; Glaser 1959; Huckfeldt and Sprague 1995; Straits 1990; Knack 1992; Kenny 1992; Beck et al 2002).

Most models of elections also make typical cognitive assumptions about information and individual rationality, in spite of the evidence from psychology that both may be severely limited (e.g. Simon 1982; Quattrone and Tversky 1988). Instead, people might use "fast and frugal heuristics" to deal with informational limitations and strategic complexities but still achieve relatively good results (cf Gigerenzer 1999; Cosmides and Tooby 1996; Lupia and McCubbins 1998).

The complexity of including all these features in a formal analytical model would overwhelm it. A closed-form solution would probably not be tractable. However, leaving some or all of these features out may yield incomplete inferences about voter and party behavior. Therefore, I develop a formal model using an alternative methodology: agent-based modeling (ABM). Like analytical models, ABMs are built on formal assumptions about agents (players in games) and how they interact. Similar to the standard analytical models, the assumptions are clearly defined, the results are stated in precise terms, and are typically easy to replicate (Gilbert and Troitzsch 1999). Unlike most analytical models, however, ABMs are usually analyzed computationally, which means they are less elegant but also less susceptible to problems of tractability. Computational models generate data to show the relationships between variables of interest. Moreover, agent-based models may make it easier to analyze paths to equilibrium, to recognize emergent patterns of interaction, and to quickly generate models like this one where interaction is especially complicated (Johnson 1998). In other words, computational modeling provides an insight into not only the outcome of a process, but the dynamics of the process itself without sacrificing the rigor of formal modeling (Nelson and Winter 2002).

In this article I describe and analyze an agent-based model of repeated elections in which voters and parties behave simultaneously. I place voters in a social context and let them interact with one another when choosing whether or not to vote. I also let parties choose the platforms they offer, and these choices may change from election to election depending on feedback from the electorate. This allows us to explore the endogenous interaction of dynamic platforms and costly turnout. In the process I relax standard assumptions of unlimited information-processing capacities and individual hyper-rationality. Citizens are limited to information they can get from their immediate neighbors. They are boundedly rational agents who use simple heuristics to make the turnout decision. Parties are assumed to be more sophisticated, optimizing their choices given their beliefs about the expected behavior of voters and their opponents. However, they form these beliefs based on limited information—they only know the results of past elections. Since agents may not have enough information, or brain power, to solve their game prospectively, I assume they are *adaptively* rational using past outcomes to make better choices in the future (cf Kollman, Miller, and Page 1992).

The computational model that I analyze generates a number of results that contribute to the interdisciplinary literature on voting. First, the average level of aggregate turnout is empirically realistic and it varies from election to election within a stable range. Second, I show that the model is consistent with much of the empirical evidence generated to test the rational calculus of voting. Turnout increases as the cost of voting decreases, the stakes of the election increase, and the margin of victory declines. Thus even though citizens have very limited information and use a very simple learning rule, they are able to respond as though they were prospectively rational to variation in the incentive to vote. Third, the model is consistent with empirical results from the literature on the social context of voting. In particular, turnout correlates highly between neighbors, and citizens who discuss politics with more neighbors are more likely to vote. Fourth, the model also generates a surprising result: when citizens are situated near people with similar preferences they are less likely to vote. In short, segregation depresses turnout. Finally, I explain why a local imitation structure inherently yields dynamics that encourage positive turnout.

The model also generates a number of results that contribute to the literature on party behavior. First, consistent with Wittman equilibrium under uncertainty, policy-motivated parties offer divergent platforms. In this setting, citizens – free to vote or abstain – serve as a source of uncertainty since the location of the median voter is changing all the time. Second, parties adjust their platforms in direct response to the vote share in the previous election. Both parties move in the direction of the previous winner and in proportion to the previous margin of victory. Third, parties are drawn not only to the median voter, but also to the median citizen since she represents the median of the pool of potential future voters. Finally, the model generates another surprising result: electorates with higher local correlation of preferences lead to a greater divergence of party platforms. This suggests that parties polarize as neighborhoods become more segregated.

In the following section I describe most the general structure and most important elements of the agent-based model of elections: how voters make their decisions and how parties choose their platforms. Then I proceed with analysis of the main results of the model: most notably, why people vote despite the cost of voting, and what electoral aspects influence party platforms. In the final section I summarize the findings and discuss application of computational models of elections in the future research.

Model

In this section I describe a simplified version of the computational model.[14] As in the standard political science model of elections, I assume that each citizen in a population has some preferred policy point on a one-dimensional left-right scale, which one can think of as liberal-conservative issue space. Two parties compete in elections, and these parties have fixed left and right preferences. The parties choose electoral platforms (see below) and each citizen chooses to vote or abstain. If a citizen turns out, she chooses the party offering the platform closest to her own preference. Votes for the left and right are counted and the election winner is determined by the majority rule. After each election, a citizen's utility is simply the negative squared distance between her preferred policy and the platform implemented by the winning party, minus the cost of voting.[15]

---

[14] Code for the R implementation can be found at http://jhfowler.ucdavis.edu

[15] $u_i = -(x_W - p_i)^2 - c$, where u_i is the utility of voter i, x_W is the platform of the winning party, p_i is the preference of voter i, and $c > 0$ is the cost of voting. If a citizen abstains, $c = 0$.

Parties are assumed to be policy-motivated: they have the same preferences and utility over the policy space as voters (a party prefers to win the election with a policy closer to its ideal point).

Parties only know their own preference point and the results of past elections. They do not know the distribution of voter preferences and, therefore, they do not know the exact location of the median voter. Moreover, some of the former voters may abstain and some of the former abstainers may vote, meaning the location of the median voter may change from election to election (Brody and Page 1973). To deal with this uncertainty I assume that parties use previous election results to learn about the voter distribution. First, they use the results of the past election to estimate the location of the median voter.[16] For example, if the left party wins in a landslide, both parties can infer that the median voter was located closer to the left platform than the right platform. Second, they use Bayesian inference to update their beliefs about the expected median voter in the coming election.[17] Given these beliefs about the electorate, the parties choose platforms by mutually optimizing their expected payoffs.[18]

Unlike parties, citizens employ a less sophisticated decision-making mechanism. I model citizens as boundedly rational agents with access to limited information. In the model, they only know the utility and turnout behavior of their immediate neighbors. This means they also do not know the true preferences of any other citizens or parties. One might argue that this is unnecessarily naive—surely people think for themselves!

---

[16] The location of the median voter m is the solution to the equation $S = \int_{-\infty}^{(x_L+x_R)/2} f(v \mid m)dv$ where S is the vote share for the left party, x_L and x_R are the party platforms, and $f(v)$ is the voter distribution (which we assume to be normal with variance 1).

[17] We assume that parties model the location of the median voter as though it were drawn from a normal distribution with unknown mean and variance. It is well known that under these conditions the expected median voter will be the sample mean of all previous observations of the location of the median voter, and the variance in the expected median voter will be the sample variance.

[18] Following Wittman (1977) we assume that the expected payoff of each party is the probability of winning times the winning payoff plus the probability of losing times the losing payoff. Parties choose a set of equilibrium policies in which neither party can achieve a higher expected payoff by changing its platform. For a detailed description of the dynamic model of policy-motivated candidates under uncertainty and Wittman political equilibrium solution see Smirnov and Fowler (2005).

However, I know from much of the empirical literature on contextual effects that local information has a powerful influence on individual voter behavior (Beck et al 2002; Fotos and Franklin 2002). Imitation has been shown to be an extremely cost-effective strategy in complex environments, even if it does not necessarily lead to the best possible outcome (e.g. Boyd and Richerson 1985).

To model local interaction I endow citizens with preferences and place them randomly on a grid.[19] I then allow them to have political discussions with other people in their neighborhood.[20] Given the constraints on information and the enormous complexity of maximizing utility over some set of future elections, citizens adopt the most successful strategies from past elections. I assume that there is an information flow among immediate neighbors with respect to the past election, in particular, whether or not they voted and how satisfied they were with the results. Since voters can learn about the turnout behavior and relative satisfaction of their neighbors they can use this information to decide whether or not to vote in the next election. Specifically, they divide people in their neighborhood between voters and abstainers, decide which type is more satisfied, and then imitate the behavior of the most satisfied group.[21]

---

[19] We assume voter preferences are independent and drawn from a standard normal distribution. Later in the article we will relax the independence assumption by assuming preferences are correlated between neighbors.

[20] We assume a Moore neighborhood structure, which means individuals typically have 8 neighbors (top, bottom, left, right, top left, top right, bottom left, and bottom right). We also assume the grid is bounded, so individuals on the edges have fewer neighbors (e.g. an individual on the left edge has five neighbors—top, bottom right, top right, and bottom right).

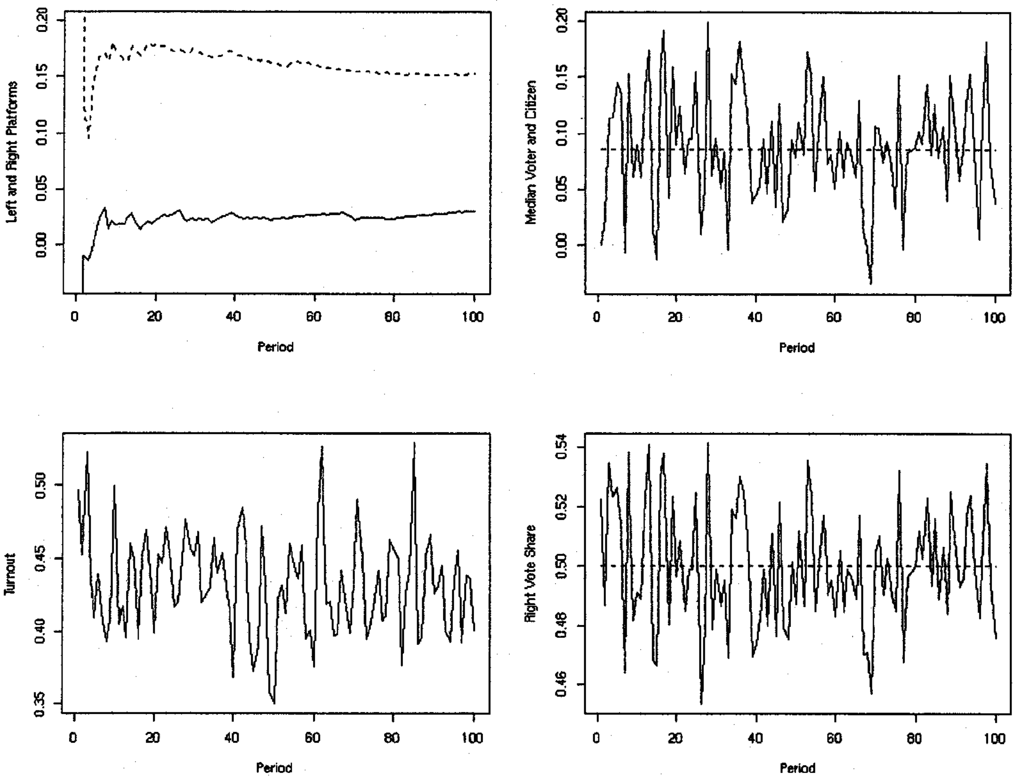
[21] There are several learning algorithms that we could choose to model this behavior, so we have deliberately chosen a simple one. Citizens have discussions with each of their neighbors and learn how satisfied they were (i.e. their utility) with the results of the previous election. Each citizen then estimates the average satisfaction with voting s^{vote} by summing the satisfaction of all voters in the neighborhood (including themselves, if applicable) and dividing by the number of voters. Similarly, they find the average satisfaction with abstaining $s^{abstain}$ by summing the satisfaction of all abstainers in the neighborhood (including themselves, if applicable) and dividing by the number of abstainers. If the number of voters or abstainers in the neighborhood is zero, then the individual repeats her action from the previous election.

Results: General Dynamics

To analyze computational results from the model I employ three strategies. First, I develop a graphical user interface (GUI) for the model so I can watch what happens to voter utilities, turnout, platforms, and other variables of interest. Computational modeling is unique in this respect because it allows us to inspect visually what is happening to the model as it progresses. This sometimes leads to hypotheses about the dynamic processes that might not otherwise have been obvious using different methodologies (Gilbert and Troitzsch 1999). Second, I produce graphs of several runs of consecutive elections. These graphs are snapshots of the dynamic behavior of one or two variables from the model and they are useful for characterizing typical boundaries and changes in the values for a given set of model assumptions. Third, I conduct multiple runs and collect data at the end of each run. This allows us to see how changes in assumptions affect how the model behaves.

In Figure 6 I present some results from a typical run of 100 elections. The lower left graph shows that turnout varies between 35% and 55%. When I let the simulation run for thousands of elections turnout never jumps out of this range: turnout seems to be significant and stable even when it is costly. The upper right graph shows how the model generates instability in the location of the median voter. Even though the preference of the median citizen remains fixed for a given run (represented by the straight horizontal line in the graph), the preference of the median voter depends on who decides to vote and changes from election to election. Notice especially that the median voter can remain to the left or right of the median citizen for several elections, indicating a period when one party's supporters are more active than the other's.

Figure 6. Results of a Single Run of 100 Elections



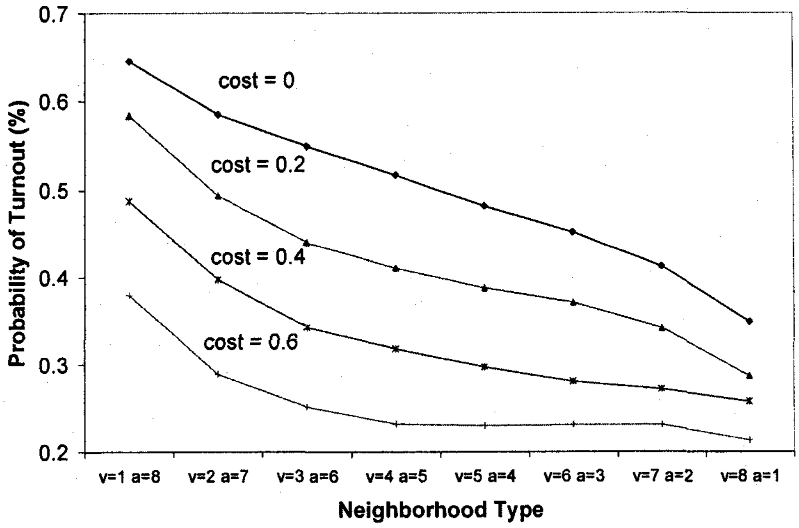
Note: For this run, I test a population of 1024 voters with independent preferences drawn from a standard normal distribution, party preferences at -1 and 1, cost of voting of 0.1, and initial probability of turnout of 0.5. In the upper left graph, solid line is the left party and dotted line is the right party. In the upper right graph solid line is the median voter and dotted line is the median citizen. In the lower right graph the dotted line marks the location of a tie (right vote share = 0.50).

The upper left graph shows how party platforms change over time to adapt to these circumstances. After a brief convergence from initial conditions and a period of instability the platforms tend to oscillate in a stable range that remains significantly far from the center. This oscillation seems to vary with the location of the median voter as parties attempt to adjust their platforms in the median voter's direction. Constant adjustment by the parties also generates variation in the margin of victory in the lower right graph as parties alternate winning and losing elections.

Why So Much Turnout?

The main source of turnout in the model has to do with imitation in a social context. I assume that citizens are boundedly-rational, acquiring information only from their neighbors in order to decide whether to vote in the next election. In the extreme cases in which everyone votes or everyone abstains, the citizen simply repeats her prior action. In other cases I can derive the expected probability that the voters in a randomly sampled neighborhood will happen to do better than the abstainers because of the random location of their preferences.[22] Figure 7 shows the probability that a randomly sampled citizen will vote given the number of her neighbors who voted in the previous election and the cost of voting.

Figure 7. Theoretical Impact of Cost and Neighborhood Type on Turnout



Note: a = number of abstainers in a neighborhood in previous election; v = number of voters in previous election (e.g. v=1 a=8 is a neighborhood with one voter and eight abstainers in the previous election. Citizens in a neighborhood like this have a 65% chance of voting if the cost of voting is 0, 58% if the cost is 0.2, 49% if the cost is 0.4, and 38% if the cost is 0.6).

---

[22] It is important to emphasize here that while citizen decisions are deterministic in the model, the distribution of preferences is stochastic. Thus utility itself is a random variable: a citizen in an n person neighborhood with v voters will vote in the next election with probability $\Pr\left(\frac{1}{v}\sum_{i=1}^{v}u^i > \frac{1}{n-v}\sum_{j=1}^{n-v}u^i\right)$.

From top to bottom, each curve in Figure 7 represents a higher cost of voting. Note that increasing the cost of voting decreases the probability of voting for all neighborhood types. This is because the cost of voting directly decreases the average satisfaction of voters in all neighborhoods. This effect is intuitive and conforms to other theoretical and empirical models. Note also that when voting is costless (the top curve), the probability of voting is about 0.5 when about half the neighborhood votes and half abstains (between 4 and 5 voters in a 9 person neighborhood). The expected utility to voters and nonvoters is the same if there is no cost to voting, so the odds that one group does better than another should be the same for both at 0.5. However, this is only true when the number of voters and abstainers is about the same. The downward slope in the curves in Figure 2 indicates that citizens with *fewer* voters in their neighborhood are *more* likely to vote and citizens with *more* voters in their neighborhood are *less* likely to vote. This suggests a negative reinforcement effect that encourages turnout. As the probability of turnout declines, so does the expected number of voters in a given neighborhood, but the probability of turnout for these neighborhoods increases as the number of voters in the neighborhood decreases.

Negative reinforcement may seem counterintuitive, but consider the fact that each citizen is essentially sampling from the population. When one sampled group is substantially larger, it is more likely to yield an average satisfaction level that is close to the population average. The smaller group is privileged because there is a better chance that they will happen to have preferences very close to the winning platform. For example, suppose that half the citizens in a neighborhood vote in the first election and voting is costly. After that, citizens decide whether or not to vote by comparing average utilities of voters and abstainers. It is likely that eventually the number of voters in the neighborhood will decrease to 1 or 2 since the cost of voting is positive. However, if one of the few remaining voters happens to have a preference that is relatively close to the platform of the winning party, the voter will be more satisfied than the abstainers. Since the number of voters in the neighborhood is small, her satisfaction will dominate the average satisfaction of turnout. As a result, her neighbors will imitate her turnout behavior. Of course, the local surge of voting will be quickly suppressed by the cost of voting and, thus, I have a local turnout-abstention cycle. The global dynamic is a combination of all the overlapping local neighborhoods, all of which experience periods of turnout and abstentions at different moments of the time. Hence a local imitation structure inherently yields dynamics that encourage turnout.[23]

The Rational Calculus of Voting

The rational calculus of voting model assumes that voters think *prospectively* about the impact of their actions on their own utility. Advocates of this model cite several empirical regularities predicted by the model as evidence that these assumptions are correct. In contrast, the model assumes that voters *adapt* to past outcomes. In Figure 8 I see that the model generates the same empirical regularities. For example, turnout is sensitive to the cost of voting. An increase from nothing to 0.1 depresses turnout by about 4 percentage points.[24] The tendency of voters to respond to higher costs with lower turnout is consistent with a broad empirical literature on the subject. For example, restrictive registration laws clearly discourage voting (Rosenstone and Wolfinger 1978; Squire et al 1987; Nagler 1991; Rhine 1995; Knack 1997, 2001; Franklin and Grier 1997; Fenster 1994; Highton 1997; Knack and White 2000; Highton 2000; Huang and Shields 2000), while liberal absentee ballot laws and all-mail elections encourage it (Oliver 1996; Karp and Banducci 2000; Southwell and Burchett 2000b).

The rational calculus of voting literature also posits that voters should be influenced by the expected benefits from voting expressed as a function of the distance between the parties and the probability of influencing the outcome of the election. The model produces both of these relationships. In the right graph of Figure 8 turnout increases with the distance

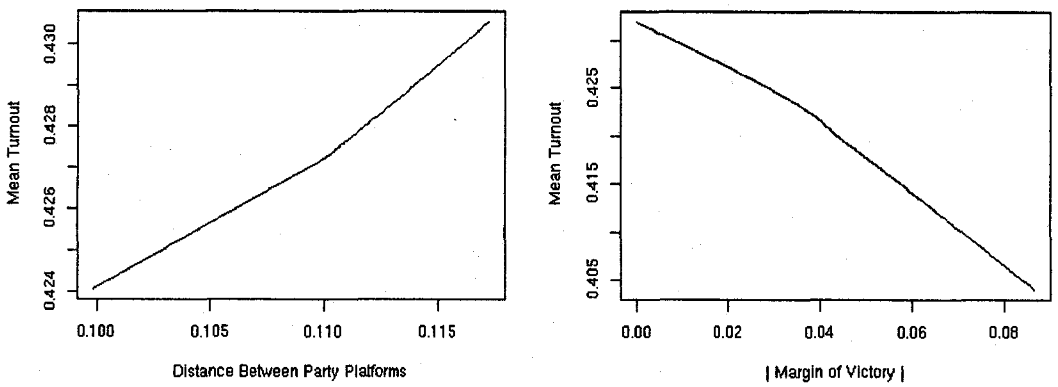
---

[23] The negative feedback mechanism not only leads to turnout when it is costly but also to abstention when it is not costly – in fact, even if we make the cost of voting negative – turnout will still be significantly less than 100% for this reason.

[24] This may seem like a trivially small cost of voting, but consider the fact that the mean distance between the left and right party platforms in our sample is about 0.2. If voters must bear a cost of 0.1 in order to vote, they are paying one half of the total benefit they would receive if they could choose the election winner. For most of the formal models cited above, the highest cost-benefit ratios that would yield positive turnout are typically several orders of magnitude smaller than this.

between party platforms, consistent with empirical work that suggests that turnout is somewhat higher in elections with higher stakes (Wolfinger and Rosenstone 1980; Boyd 1989; Hansen, Palfrey, and Rosenthal 1987; Forgette and Sala 1999; Jackson 2000) and a larger distance between the parties (Kaempfer and Lowenberg 1993).

Figure 8. Determinants of Voter Behavior



Note: Each graph based on 1000 simulations of a population of 1024 voters with independent preferences drawn from a standard normal distribution, party preferences at -1 and 1, and initial probability of turnout of 0.5. Upper right and lower graph assume a cost of voting of 0.1. Cost of turnout was varied from 0 to 1 in the upper left graph.

In the right graph turnout varies inversely with the closeness of the election: participation decreases as the margin of victory by one of the parties increases. This effect is consistent with an empirical literature that has tried to use the closeness of an election as a proxy for how voters perceive the likelihood of affecting the outcome. Though the relationship has been questioned by some (Key 1949; Matsusaka 1993; Ansolabehere and Iyengar 1994; Kirchgassner and Himmern 1997; Kunce 2001), the weight of the evidence seems to point to a small but significant correlation between closeness and turnout (Cox and Munger 1989; Berch 1993; Jackson 1983; Hanks and Grofman 1998; Grofman et al 1998; Nalebuff and Shachar 1999; Alvarez and Nagler 2000).

The fact that the model produces results consistent with the rational calculus of voting suggests that the adaptation model for citizens is sufficiently sophisticated that they are able to learn to vote more often when it would make them better off—that is, when costs are low, stakes are high, and elections are close. However, turnout is still quite high relative to a model in which citizens are perfectly informed and strictly utility-maximizing. To see if this discrepancy is associated with limited information, I alter the model slightly by endowing citizens with memory.

Memory permits citizens to combine information from previous elections with new information about the merits of voting and not voting. Specifically, a memory parameter governs how new information is weighed relative to previous information.[25] If this parameter is set to zero, then citizens only remember the results of the past election. As the parameter increases towards one, they remember more and more of the past and as a consequence the relevance of the current election decreases. The graph in Figure 9 shows the effect of increasing citizen memory. As voters acquire more information about the relative merits of voting and abstaining, they choose to abstain in greater numbers.

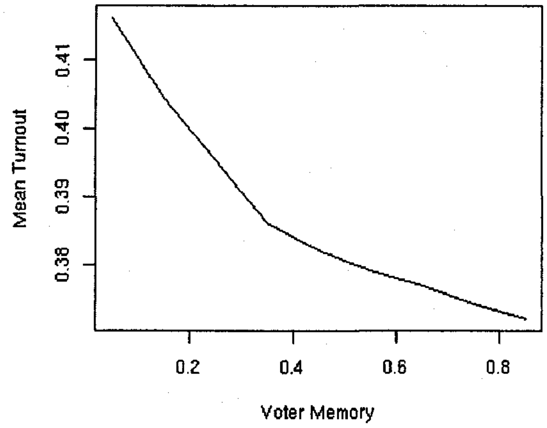
The negative relationship between memory and turnout suggests that limited information about the costs and benefits of voting plays an important role in supporting high levels of participation. To make sense of this, think of the extreme case. Without memory, the only information citizens have is the relative satisfaction levels of their neighbors and themselves for the most recent election. With memory, citizens have access to all this information, plus some of the information they acquired in previous elections. As memories lengthen, the number of individual satisfaction levels that go into the average satisfaction level increases, improving the estimate of the relative costs and benefits of participation.

---

[25] Let M be a memory parameter. As above, citizens find the average satisfaction level of voting and abstaining for the current election, but they now weight the new information with previous estimates of the average satisfaction levels for voting and abstaining:

$$s_{t+1}^{vote} = (1-M)\frac{1}{v}\sum_{i=1}^{v} u_t^i + Ms_t^i,\ s_{t+1}^{abstain} = (1-M)\frac{1}{a}\sum_{i=1}^{a} u_t^i + Ms_t^i$$

Figure 9. Effect of Memory on Turnout



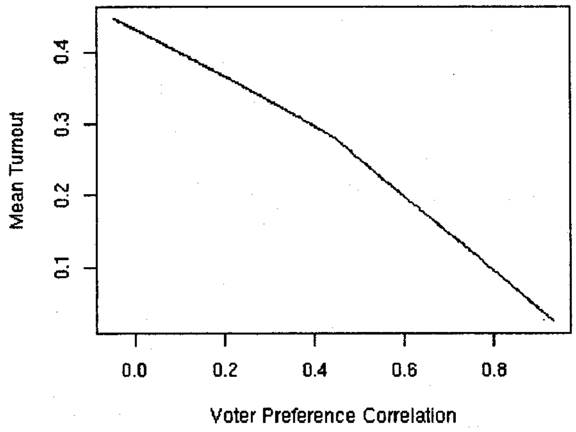
Note: Based on 1000 simulations of a population of 1024 voters with independent preferences drawn from a standard normal distribution, party preferences at -1 and 1, and initial probability of turnout of 0.5. Voter memory was varied from 0 (least weight on past information) to 0.9 (most weight on past information).

Social Networks and Turnout

The model produces results that are consistent with findings related to social networks. At the level of the individual voter I find correlation in vote strategies between neighbors. For the baseline simulation, this correlation is about $\rho = 0.29$ and it does not change much when I try different combinations of parameters. This result conforms to the finding that turnout is correlated between friends, family, and co-workers (Lazarsfeld et al 1948; Berelson et al 1954; Campbell et al 1954; Glaser 1959; Huckfeldt and Sprague 1995; Straits 1990; Knack 1992; Kenny 1992, 1993). One might argue that this is a trivial result. After all, the model assumes that voters imitate their neighbors, so I should expect to find some correlation in turnout behavior. However, I emphasize that this is the only theoretical model I are aware of that generates correlated turnout. What it suggests is that models that do not embed their citizens in a social network context may be omitting an important feature of the real world that is relevant to turnout behavior.

The social network context I have supposed so far is artificial in a very important way. I assume that individual preferences are not correlated. The probability of a liberal speaking to another liberal in the model is the same as the probability of a liberal speaking to a conservative. However, a consistent finding in the social voting literature is that people tend to segregate themselves into like-minded groups. As a result, most social ties are between people who share the same interests. Even when people with ideological or class-based interests are not surrounded by like-minded individuals in their *physical* neighborhoods and workplaces they tend to withdraw and form relationships *outside* those environments (Huckfeldt, Johnson, and Sprague 2003; Huckfeldt and Sprague 1987, 1988; Noelle-Neumann 1984; Gans 1967; Berger 1960). Thus preferences between acquaintances tend to be highly correlated. For example, in the Indianapolis-St. Louis Election Study the correlation in liberal-conservative ideology is ρ=0.66, while the correlation in party preference is ρ=0.54.

Figure 10. The Effect of Preference Correlation on Turnout



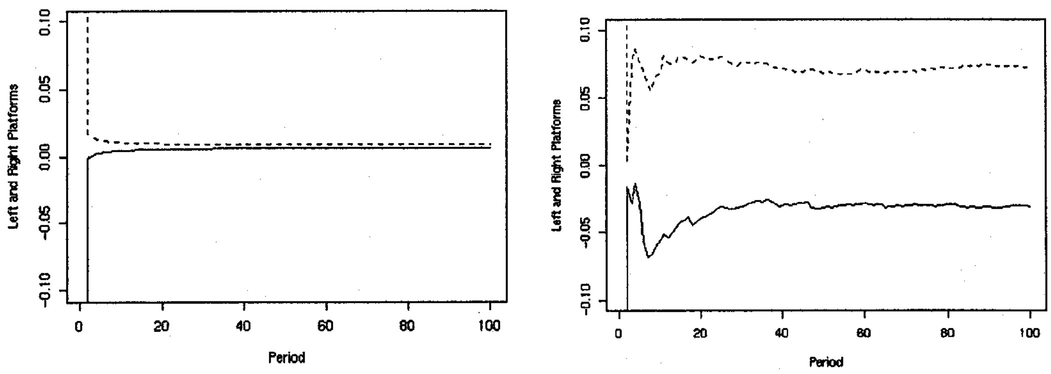
Note: Based on 1000 simulations of a population of 1024 voters with independent preferences drawn from a standard normal distribution, party preferences at -1 and 1, cost of voting of 0.1, and initial probability of turnout of 0.5. Preference correlation was varied from 0 to 0.95.

What effect does the concentration of shared interests have on the model? Figure 10 shows that preference correlation has a dramatically negative effect on turnout. When a citizen has discussions with a diverse group, it is more difficult to discern the costs and benefits of voting. However, when all a citizen's neighbors are just like her, she is more likely to free ride. To see why, suppose an extreme case in which everyone in a citizen's neighborhood has the same preference. When comparing the average satisfaction level of voters and abstainers, the benefits will be exactly the same for everyone. The only thing that differentiates the voters from the abstainers is the cost of voting. Thus, it will be easy to figure out that free riding makes sense. Now suppose the opposite case in which neighbors have heterogeneous preferences. Even though all voters pay a cost of voting, some voters will be very satisfied because they happen to be located close to the winning candidate. Conversely, even though abstainers do not pay a cost of voting, some will be very dissatisfied because they have preferences that are far away from the winning candidate. Thus, as preference correlation decreases, the relationship between satisfaction level and turnout behavior breaks down and it becomes more difficult to discern the advantage of free riding. In short, social segregation hurts participation.

Party Behavior

Turning to party behavior, I note that the model generates a substantial degree of platform divergence (see Figure 6 above). The game-theoretic literature suggests that uncertainty is a necessary condition for platform divergence (Wittman 1977; Calvert 1985). These models introduce an exogenous source of uncertainty, but in the model uncertainty is generated endogenously by variation in voter turnout. The location of the median voter changes from election to election as new sets of voters show up to the polls. Figure 11 compares results when I fix voter turnout to those when I allow it to vary.

Figure 11. Effect of Fixed and Variable Turnout on Party Behavior

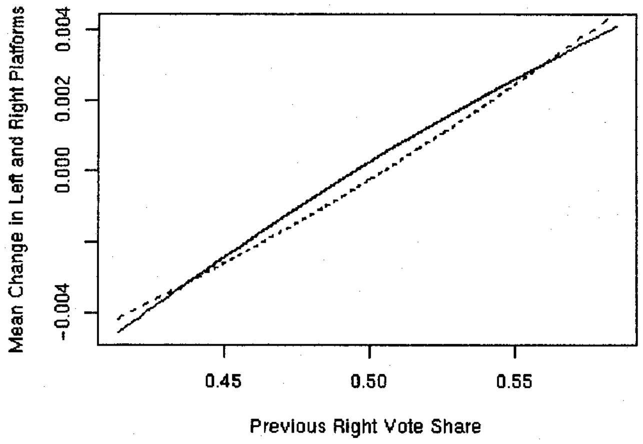


Note: Example based on a population of 1024 voters with independent preferences drawn from a standard normal distribution, party preferences at -1 and 1, and cost of voting of 0.1. Solid line is the left party and dotted line is the right party. The left graph assumes fixed voter turnout and the right graph assumes variable voter turnout.

When I fix turnout and the location of the median voter is constant, the parties quickly infer its location and converge. When I allow voters to choose whether or not to vote, the platforms diverge. Clearly, parties behave differently when turnout behavior is allowed to vary, suggesting that it may be important to model both voters and parties simultaneously as I do here.

The model also suggests that platform divergence may result from parties choosing strategies that react positively to the margin of victory. Figure 12 shows that both parties typically move their platforms in the direction of the winning candidate and in proportion to the margin of victory.

Figure 12. Effect of Vote Share on Party Behavior

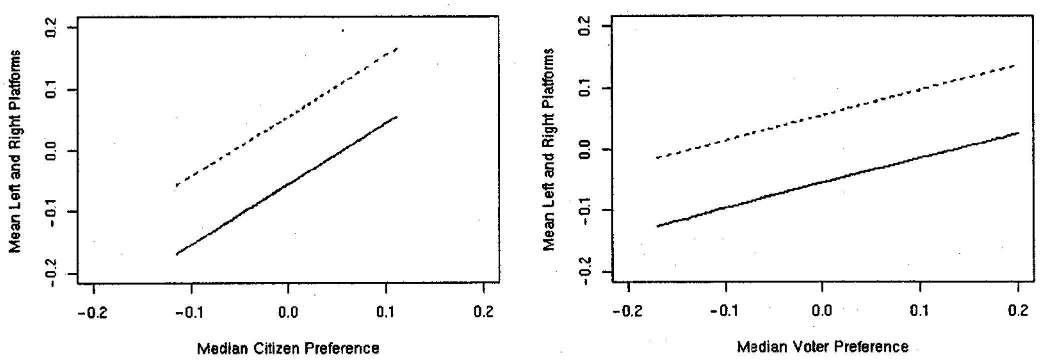


Note: Based on 1000 simulations of a population of 1024 voters with independent preferences drawn from a standard normal distribution, party preferences at -1 and 1, cost of voting of 0.1, and initial probability of turnout of 0.5. Solid line is the left party and dotted line is the right party.

For example, if the left wins a close election, both parties will shift slightly to the left. If the left wins in a landslide, both parties will shift a lot to the left. This is because a landslide victory causes the winning party to infer that it can win with a platform that is closer to its own preferences. It also causes the losing party to learn that it must moderate in order to be competitive in the next election. The relationship between platforms and vote share is consistent with the literature on Presidential mandates (Conley 2002; Kingdon 1966), a more detailed analysis of Wittman equilibrium (Smirnov and Fowler 2003), and recent evidence that shows past vote share affects the ideology of US Senate candidates (Fowler 2003).

The effect of these strategic interactions is that parties try to adapt to the (unknown) positions of the median voter and the median citizen. In Figure 6 I showed that in a given run the median voter changes frequently while the median citizen remains constant. Parties have a short-term incentive to exploit the former if there tends to be some persistence in the set of voters who turnout from one election to the next. However, they also have a long-term incentive to stay close to the median citizen since this represents the pool of all possible voters in future elections. Figure 13 shows that platforms tend to track *both* changes in the location of the median voter and the fixed location of the median citizen. Interestingly, the parties are more sensitive to the location of the median citizen than the median voter, which implies that parties pay more attention to the long-term shape of the electorate rather than the short-term changes.

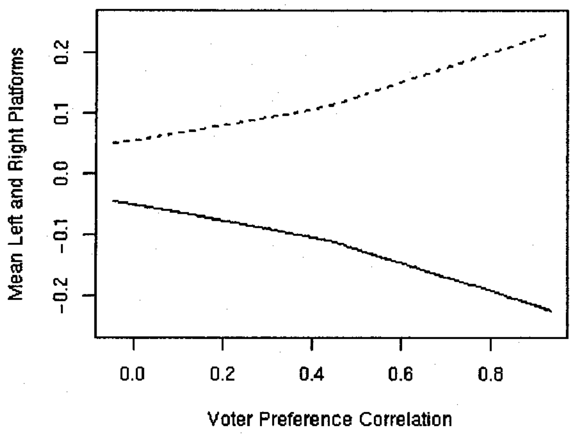
Figure 13. The Effect of Median Citizen and Median Voter Preferences on Party Behavior



Note: Based on 1000 simulations of a population of 1024 voters with independent preferences drawn from a standard normal distribution, party preferences at -1 and 1, cost of voting of 0.1, and initial probability of turnout of 0.5. Solid line is the left party and dotted line is the right party.

Finally, I highlight a surprising interaction between parties and voters. Figure 14 shows that increasing preference correlation among voters dramatically increases platform divergence. This is because preference correlation tends to increase variance in the vote share. Heterogeneous neighborhoods will have one or two citizens switching their behavior when the parties adjust slightly to the left or right, but homogeneous neighborhoods will have several citizens switching together—small changes in the location of the parties can quickly lead to waves of imitation among supporters of one of the parties. Whole neighborhoods teeter on the brink of voting or not and the result is to increase swings in electoral outcomes. This increases uncertainty about the location of the median voter and has a corresponding effect on the parties. In short, self-segregation yields party polarization.

Figure 14. The Effect of Preference Correlation on Party Behavior



Note: Based on 1000 simulations of a population of 1024 voters with correlated preferences drawn from a standard normal distribution, party preferences at -1 and 1, cost of voting of 0.1, and initial probability of turnout of 0.5. Voter preference correlation was varied from 0 to 0.95. Solid line is the left party and dotted line is the right party.

Heterogeneous neighborhoods will have one or two citizens switching their behavior when the parties adjust slightly to the left or right, but homogeneous neighborhoods will have several citizens switching together—small changes in the location of the parties can quickly lead to waves of imitation among supporters of one of the parties. Whole neighborhoods teeter on the brink of voting or not and the result is to increase swings in electoral outcomes. This increases uncertainty about the location of the median voter and has a corresponding effect on the parties. In short, self-segregation yields party polarization.

Summary

The subject of elections, including turnout and platforms dynamics, is challenging for all social scientists. One of the main reasons for this difficulty lies in the fact that various elements of the electoral process are easier to study separately. I believe that an interdisciplinary approach, built upon contributions from several social science disciplines, will lead us to a better understanding of the subject. The agent-based model I propose is built upon a number of important contributions by sociologists (social context of voters), psychologists (bounded rationality and use of heuristics), economists (platforms dynamics and turnout decision), anthropologists (cultural influence exemplified by imitation), and last, but not least, political scientists (interdependence of voters and candidates, dynamic nature of the electoral competition, empirical analysis of observations).

The model yields several findings consistent with the empirical literature on parties and voters and suggests some relationships that have not yet been tested (see Table 13 for a summary). The central result is that turnout is significant, platforms diverge, and they both vary over time in an empirically realistic way. These phenomena emerge when I allow both turnout and platform strategies to adapt to one another over time. Making citizens boundedly rational and placing them in a social context turns out to be important. A closer looks at the model neighborhoods shows that local imitation in a social network inherently yields negative feedback dynamics that encourage turnout. The effect is further amplified by the natural limits on the information-processing capacities of the citizens such as the length of memory. On the other hand, local correlation of preferences appears to decrease individual propensity to turn out, which implies that ideologically homogenous communities are least likely to vote. The model also conforms to findings from the social voting literature. Citizens appear to be affected by the turnout decisions of their neighbors.

Turning to parties, the model yields several empirical implications. Allowing turnout to vary endogenously generates uncertainty about the location of the median voter and causes party platforms to diverge. I also note that parties pay attention to electoral mandates as they try to estimate the location of the median voter to remain competitive. This could help to explain empirical work that shows the ideology of US Senate candidates and expectations of economic policy are sensitive to previous vote share (Fowler 2002; Fowler 2003). The model also shows that party platforms tend to correlate with changes in the position of both the median voter and the median citizen, with parties being more sensitive to the latter. Finally, a higher degree of local preference correlation among voters leads to greater platform divergence. Voter segregation yields party polarization.

In conclusion, though the model generates relationships that correspond to much of what we know about turnout and platforms, it is important not to read too much into the results. There are many factors that I have not included here that may affect turnout and platforms such as socioeconomic status, endogenous voter and party preferences, multidimensional issue space, multiple parties, multiple districts, different electoral institutions, political institutions like legislatures, and so on.

Methodologically, elections may be seen as an unlikely subject for application of evolutionary models. Indeed, the choice of platforms and turnout decision are the behaviors which emerged only very recently and no corresponding innate propensities can exist. Nevertheless, the evolutionary models of adaptive learning and evolutionary imitation appear to be an insightful modeling analogy, a metaphor that may help explain very puzzling behavior. Given these models, I assume that politicians are retrospectively rational, choosing electoral platforms given the past electoral history. Further, I assume that voters use very simple evolutionary heuristic – an informational shortcut – which is "observe others, copy the most successful behavior." These evolutionary assumptions lead to the behavior which is consistent with the empirical evidence. We observe people voting in large numbers. We also observe political parties offering divergent platforms. The classical rational choice theory, on the other hand, predicts no turnout and the convergence of politicians' platforms.

Due to emerging complexity, the electoral model presented in this chapter was computational. I will return to the method of evolutionary computer simulation in chapter 7 devoted to multilevel selection. It is important to note that whenever possible analytical models are preferred to the computational models. Because the parameter space is infinite we can never be absolutely sure if the results of a computer simulation are general (see chapter 7 for further discussions and ways to improve generality). In the next chapter, I turn to evolutionary game theory to show how an analytical approach can provide insights into a challenging evolutionary puzzle of human cooperation.

Table 13. Electoral Model: Summary of Results.

Result	Consistent with
Turnout is significant and stable	Mackie & Rose 1997
Turnout depends negatively on voting costs	Rosenstone & Wolfinger 1978; Squire et al 1987; Nagler 1991; Rhine 1995; Knack 1997, 2001; Franklin & Grier 1997; Fenster 1994; Highton 1997; Knack & White 2000; Highton 2000; Huang & Shields 2000; Oliver 1996; Karp & Banducci 2000; Southwell & Burchett 2000b
Party divergence increases turnout	Wolfinger & Rosenstone 1980; Boyd 1989; Hansen et al 1987; Jackson 2000; Kaempfer & Lowenberg 1993
Turnout increases with the closeness of the election	Cox & Munger 1989; Berch 1993; Jackson 1983; Hanks & Grofman 1998; Grofman et al 1998; Nalebuff & Shachar 1999; Alvarez & Nagler 2000
Longer voter memories reduce turnout	*Original result*
Decision to vote depends on turnout behavior of socially-connected peers	Lazarsfeld et al 1948; Berelson et al 1954; Campbell et al 1954; Glaser 1959; Huckfeldt & Sprague 1995; Straits 1990; Knack 1992; Kenny 1992, 1993
Local imitation yields positive feedback for turnout	*Original result*
Ideological segregation reduces turnout	*Original result*
Parties diverge	Peltzman 1984; Grofman et al 1990; Jung et al 1991; Poole & Rosenthal 1984; Hansson & Stuart 1984; Lindbeck & Weibull 1993; Wittman 1977
Parties respond to past margins of victory	Conley 2001; Kingdon 1966; Fowler 2002; Smirnov and Fowler 2005
Parties respond *both* to median voter and median citizen	*Original result*
Ideological segregation yields polarized parties	*Original result*

CHAPTER SIX

FROM ANALOGY TO HOMOLOGY: THE EVOLUTION OF ALTRUISTIC PUNISHMENT

In this chapter, I use evolutionary game theory to examine evolutionary stability of altruistic punishment as a potential solution to the puzzle of human cooperation. The phenomenon also has a large number of theoretical links to political science, economics, psychology, and evolutionary biology. Using altruistic punishment as explanatory variable, the chapter makes an attempt to integrate the study of cooperation and the problem of collective action in social sciences with the research on evolutionary adaptedness of altruistic behavior in life sciences.

Classical evolutionary game theory is based upon two concepts: evolutionary stable strategy (ESS) and replicator dynamics. ESS is a useful refinement of the Nash equilibrium solution (each ESS is necessary an NE while not all NE are ESS). However, ESS does not tell us a full story since it does not describe the dynamics of the convergence to a steady state. This is where the second concept – replicator dynamics – becomes useful since it tells us exactly how properties and composition of the population changes over time. In this chapter I apply both concepts of evolutionary game theory to examine robustness of altruistic punishment as a type. First, using the concept of ESS, I examine whether altruistic punishment is a viable behavior that could have survived an evolutionary process. Second, using the analysis of replicator dynamics I examine what can happen *on the way* to an equilibrium (steady states). In this respect, replicator dynamics tell us a *story* how a community could have evolved depending on the parameters of the game and its starting conditions.

The analysis also has important substantive implications. Recognizing altruistic punishment as an important factor in successful management of the commons has immediate policy impact. One such implication is that to prevent the tragedy of the commons we do not always have to get rid of the commons by means of privatization or centralized coercion. Instead we may want to re-create and reinforce the commons as a common-property regime with its own norms and a certain degree of independence.

Altruistic behavior provides benefits to others while the altruists incur a cost (Hamilton 1964; Trivers 1971). Existence of such behavior is puzzling for scholars from social and life sciences. If people are rational, why do they choose A over B when A provides higher individual utility? More importantly, how could such behavior have possibly survived in the evolutionary process? Formal models of altruistic and selfish behavior, most notably the Prisoners' Dilemma game and the model of the commons (Hardin 1968; Olson 1965), increased the interest in the paradox even further and led to an unprecedented interdisciplinary research. The models predict that rational behavior will lead to mutual defection, over-exploitation of resources, and, ultimately, demise into the Hobbesian world when the life of man is "solitary, poor, nasty, brutish, and short" (Hobbes, *Leviathan*). Strict dominance of the selfish behavior made Hardin pessimistically declare that there is no technical solution to the tragedy of the commons (Hardin 1968).

The chapter pursues two goals. First, it introduces altruistic punishment as a viable technical solution to the problem of collective action. Second, using altruistic punishment as explanatory variable, the paper makes an attempt to integrate the study of cooperation and the problem of collective action in social sciences with the research on evolutionary adaptedness of altruistic behavior in life sciences. The discussion of altruistic punishment as a technical solution is based upon simple evolutionary and game theoretic models. Parameters of the models, underlying assumptions, and the conditions necessary for the maintenance of cooperation are presented through the prism of rational choice theory, evolutionary psychology, and evolutionary game theory. In addition to the discussion of theoretical aspects of the problem, the paper has references to the relevant experimental research and empirical examples, which suggests that the theoretical interest in the phenomenon of altruistic punishment is shared not only across disciplines but also different methods: theoretical models, laboratory work, and field research.

Depending on the model, altruistic punishment (AP) is a strategy, type, or behavior, which can be described as follows. In the Prisoners' Dilemma AP is "Cooperate. If the other player defects, punish[26]. Cooperate." In terms of collective action AP can be described as "Contribute to the public good. Punish those who did not. Continue contributing to the public good." In the commons AP is the following behavior: "Use the available resources so as to maximize the communal utility. Punish those who over-use, or over-exploit, the resources. Continue using resources as in step one." Existence of altruistic punishment as an individual psychological response, a communal norm, or an element of institutional design, may add to our understanding why human behavior is characterized by both cooperation and defection, altruism and selfishness.

Experimental research on cooperation provides ambiguous results. In laboratory setting, subjects choose both cooperation and defection (Plott 1983; Isaac, Walker, and Thomas 1984; Orbell, Van de Kragt, and Dawes 1986; Ledyard 1995; Lubell and Scholz 2001). Similarly, field research has identified two large classes of equilibria in the evolution of commons: successful management of the common-pool resources and the tragedy of the commons (Ostrom 1990; Ostrom, Gardner, and Walker 1994). These facts add fuel to the theoretical puzzle of human cooperation. In social sciences, the emphasis has been on game theoretic and computational models of iterated n-person Prisoners' Dilemma (Axelrod 1980; Kreps et al 1982; Fudenberg and Maskin 1986; Binmore 1994). In this context, cooperation can be a product of rational individual choice. In life sciences, scholars have been more concerned with the conditions under which cooperation can be adaptive in the evolutionary process. In this case, altruism can be thought of as a genetic trait that individuals inherit according to their fitness differential (Hamilton 1964, 1975; Price 1972; Trivers 1971; Alexander 1987; Dugatkin 1997; Sober and Wilson 1998). The puzzle of human cooperation has also lead to an unprecedented collaboration between social and life sciences and much of the research nowadays takes place across disciplines. Notable examples of such collaboration are the models of evolutionary game theory (Smith

---

[26] To be exact, "punish such that the cost of free-riding minus the cost of punishment would be less than the value of mutual cooperation." Otherwise, the punishment is useless.

1982; Bendor and Swistak 1997), models of the cultural transition (Boyd and Richerson 1985, 1992), and a variety of agent-based models of the population dynamics (Axelrod 1997; Boyd et al 2003; Orbell et al 2003).

These interdisciplinary advances, however, had a limit effect on the actual policies, management of the commons, and the problem of collective action. Two main policy prescriptions how to avoid the tragedy of the commons remain (1) privatization of the common-pool resources (Smith 1981; Anderson and McChesney 2003), and (2) centralized coercion by the government (Hardin 1968; Ophuls 1973). Both policies presume that cooperation is hard to achieve or plain impossible. A possibility of decentralized self-enforcement of cooperation is quickly dismissed as not viable: for an individual the costs of monitoring and enforcement can be very high while the benefits are divided among members of the commons. In fact, policing is a public good, which merely shifts the problem of collective action one step further (Elster 1989). As a result, the study of altruism has focused on [altruistic] cooperation, and ignored, until recently, altruistic punishment as a redundant problem. Close examination of the phenomenon, however, reveals that there are important asymmetries between cooperation and punishment. If individuals are willing to incur the cost of punishment, I observe evolution of altruistic behavior which is more robust than the traditional non-punishing cooperation (cf Boyd et al 2003; Boyd and Richerson 1992; Henrich and Boyd 2001). Hence, the first question that I have to address is whether, in fact, individuals are willing to punish defectors even when such behavior is irrational in economic sense.

According to recent experimental evidence, the punishment of defectors is indeed a wide-spread phenomenon (Fehr and Gachter 2002). Cooperators are often willing to punish defectors even when the punishment is costly, when the game is not repeated between the same players, and when reputation-building is excluded, i.e., other players do not observe the act of punishment. Such punishment is called altruistic since it imposes a cost on the individual while not producing any material benefit. Similar results have been achieved by scholars representing different disciplines: political science, psychology and evolutionary psychology, behavioral and experimental economics (Orbell, Van de Kragt, and Dawes

1986; Yamagishi 1986; Ostrom, Walker, and Gardner 1994; Price, Tooby, and Cosmides 2002; Fehr and Gachter 2000, 2002). The work by Fehr and Gachter turned out to be especially important since the design of the experiments excluded a possibility that the punishment was non-altruistic, which is a type of punishment used in most game theoretic models explaining cooperation, e.g., trigger strategies. A possible explanation of altruistic punishment is that emotions, such as anger, are responsible for the seemingly irrational behavior. The explanation, however, begs a question: if I have emotional responses to defection, where does this propensity come from? If altruistic punishment decreases individual utility, how could such emotions and behavior have possibly evolved?

The phenomenon of altruistic punishment would not deserve much attention if laboratory results were the only evidence for it. This is not the case as field studies support the experimental research. Various forms of costly self-enforcement of cooperative behavior appear to be a wide-spread custom in communities around the world. From fisheries (Acheson 1975; Dyer and McGoodwin 1994; Crean and Symes 1996; Leal 1998; Berkes et al 2001) to irrigation systems (Tang 1992; Ostrom and Gardner 1993; Mabry 1996) to grazing lands (Netting 1981; Ellickson 1991; Anderson 1995) to forests and wildlife (Bromley 1992; Kibreab 2002), decentralized punishment of free-riding and overexploitation is a regular institutional arrangement devised to discourage opportunistic behavior (see also Ostrom 1990 and Ostrom, Gardner, and Walker 1994 for a comprehensive overview of self-governing commons).

What is so special about altruistic punishment that makes it a widespread communal custom as reflected in the field evidence, and a strong individual propensity as reflected in the experimental research? I believe there are three reasons behind the empirical evidence for altruistic punishment: economic, evolutionary, and normative (a more detailed and formal discussion of the respective reasons is presented in the second and third sections of the paper). Altruistic punishment is economically rational if pre-commitment is possible (cf chain store paradox as in Kreps and Wilson 1982). A rational individual will *have* to cooperate if the other player is pre-committed to the irrational action: punishment of defection. As a result both players are better off since no one has an incentive to free-ride[27]. The evolutionary reason behind altruistic punishment is its self-amplifying character. Unlikely cooperation, the success of altruistic punishment is frequency dependent: as the proportion of AP strategies increases in the population, defectors receive lower payoff. In a community where altruistic punishment becomes a wide-spread norm, defection disappears altogether[28]. To the contrary, a community characterized by the non-punishing cooperation is most vulnerable to defection.

The third reason behind altruistic punishment is normative. Altruistic punishment as a mechanism that preserves cooperation is arguably more preferred to Tit-For-Tat, a retaliatory mechanism that entails punishment by defection. Although theoretically Tit-For-Tat can be an effective strategy leading to cooperation (Axelrod and Hamilton 1981), it is often inappropriate and ineffective in the real world. First, punishing other defectors by means of one's own defection also harms individuals who cooperate. Second, one's own defection in the presence of other Tit-For-Tat strategies leads to defection by others. In contrast, under altruistic punishment, an individual always cooperates for the benefit of other cooperators and punishes only those who deserve it. Field studies confirm that individual violations of rules lead to punishment of those who are responsible, instead of leading to cascading defections by the rest of the group (Ostrom 1990; Ellickson 1991; Anderson 1995). Overuse of a communal irrigation system leads to the punishment of the responsible individual, and not the increase of consumption by the rest of the community (Tang 1992; Mabry 1996). Similarly, underprovision of public good such as putting efforts into a buffalo hunt leads to the punishment of shirkers instead of shirking by the rest of the group (Anderson, personal communication).

In addition to experimental and field support, the phenomenon of altruistic punishment has strong theoretical links across disciplines. In political science and economics such links include the theory of strong reciprocity (Sethi and Somanathan 1996; Gintis 2000), the models of quasi-voluntary compliance (Levi 1988), the Norms game (Axelrod 1986), punishment by exit (Vanberg and Congleton 1992), and some of the

---

[27] See Section II for the discussion of psychological and cultural mechanisms that make pre-commitment possible.

[28] See Sections III for the discussion of evolutionary stability of altruistic punishment.

literature on sanctions (Romer 1984; Shavell 1987; Nossal 1989). Examples from psychology and anthropology are the evolutionary models of punishment (Boyd and Richerson 1992) and the theory of punitive sentiments (Price, Cosmides, and Tooby 2002). Finally in life sciences, altruistic punishment is invoked by the notion of moralistic aggression (Trivers 1971), examples of punishment in animal societies, or negative reciprocity (Clutton-Brock and Parker 1995), mutual policing (Frank 1995), and repression of selfishness in the context of group selection (Sober and Wilson 1998).

Theoretical, experimental, and field research which is directly related to the phenomenon of altruistic punishment is summarized in Table 1[29]. The impressive interdisciplinary research, some of which represented in the table, sets a stage for the presentation of altruistic punishment in the context of the Prisoners' Dilemma game, the discussion of cognitive, cultural, and institutional conditions for its success, and the analysis of AP as a trait in the evolutionary process under individual and group selection pressures.

---

[29] The works included in the table by no means represent a complete picture of the research on human cooperation related to the idea of altruistic punishment but the selection should give the reader some idea about the scope of the interdisciplinary interest in the question. Moreover, some of the relevant literature on pre-commitment, reputation building, and evolutionary game theory is not in the table as less directly related.

Table 14. Altruistic punishment across different methods and disciplines.

Theory	Experimental Research	Field Studies
Political Science - The Norms game (Axelrod 1986) - Quasi-voluntary compliance (Levi 1988) - Punishment by exit (Vanberg and Congleton 1992) ***Political Economy*** - Sanctions (Nossal 1989; Shavell 1987; Romer 1984) ***Economics*** - Strong reciprocity (Gintis 2000) - Formal models of punishment of defectors (Binmore 1994; Sethi and Somanathan 1996) ***Psychology/Anthropology*** - Evolutionary model of retribution (Boyd and Richerson 1992) - Punishment of the second-order free-riders (Henrich and Boyd 2001) ***Biology*** - Reciprocal altruism and moralistic aggression (Trivers 1971) - Negative reciprocity and punishment in animal societies (Clutton-Brock and Parker 1995) - Mutual policing and repression of competition (Frank 1995) ***Interdisciplinary*** Computer simulation of group selection-based evolution of altruistic punishment (Boyd et al 2002)	***Political Science*** - Enforcement of cooperation (Orbell, Van de Kragt, and Dawes 1986) - Punishment of defectors for a fee (Ostrom, Walker, and Gardner 1994) ***Experimental/Behavioral Economics*** - Altruistic punishment (Fehr and Gachter 2002) - Effects of sanctions on human altruism (Fehr and Rockenbach 2003) - Punishment by rejection in the Ultimatum games (Camerer and Thaler 1995) ***Psychology/Evolutionary Psychology*** - Provision of a sanctioning system as a public good (Yamagishi 1986) - Effect of the probability of punishment (Kurzban et al 2001) - Punitive sentiments (Price, Tooby, and Cosmides 2002) ***Interdisciplinary*** - Cooperation, reciprocity, and punishment in fifteen small-scale societies (Henrich et al 2001)	- Self-governance of the commons (Ostrom 1990; Ostrom, Gardner, and Walker 1994; Bromley et al 1992; Alston, Libecap, and Mueller 1999; Berkes et al 1989) - Managing small-scale fisheries (Leal 1998; Berkes et al 2001; Crean and Symes 1996) - Self-governing irrigation systems (Tang 1992; Ostrom and Gardner 1993; Mabry 1996) ***Specific cases:*** - New Zealand fisheries (Arbuckle and Metzger 2000) - Irrigation in medieval Valencia (Glick 1970) - Agroforestry by Mobisquads in Ghana (Veit et al 1991) - American Indian societies (Anderson 1995) - Order without law in cattlemen ranges, Shasta county, CA (Ellickson 1991) - California gold rush (Umbeck 1981) - Dars of Sudan (Kibreab 2002) - Spanish huertas (Maass and Anderson 1978) - Timber forest management in Nepal and Japan (Sakurai et al 2001) - Trochus management in Indonesia (Ruttan 1998)

The empirical importance of costly self-enforcement is not reflected in the Prisoners' Dilemma game, which has been a standard model predicting the tragedy of the commons[30]. Although the model captures the crux of the problem of collective action, it is possible that certain fundamental aspects of real world dilemmas, such as punishment of defectors, are missing. Even the prisoners in the story are likely to be aware of potential retributions typical in the criminal world for not keeping silent. A simple extension of the standard PD game illustrates potential importance of altruistic punishment (see Figure 1).

Table 15. Prisoners' Dilemma Game with a Costly Punishment Option.

	Cooperator (p)	Altruistic Punisher (q)	Defector (1-p-q)
Cooperator	*1, 1*	*1, 1*	*0, k*
Altruistic Punisher	*1, 1*	*1, 1*	*-Y, k-zY*
Defector	*k, 0*	*k-zY, -Y*	*m, m*

Note: $k>1$, $m<1$, $z>0$, $Y\equiv\frac{X}{C}>0$, where X is the cost of punishment to the punisher and C is the benefit from mutual cooperation.

Altruistic punisher is a cooperator that sacrifices some of his utility in order to punish the defector, thus, decreasing his utility by zY. In this game, mutual defection remains a Nash equilibrium: no player has an incentive to deviate from his strategy. However, altruistic punishment is the other Nash equilibrium if the cost of punishment for the defector is greater than the benefit from free-riding, $1>k-zY$. The cooperative equilibrium gives both players higher utility than the mutual defection equilibrium, which makes it a salience equilibrium. At the same time, altruistic punishment is weakly dominated by cooperation: playing cooperation is always at least as good as playing

---

[30] In the two-player version of the game, each player (prisoner) has two available strategies: cooperate (maintain silence) and defect (confess). Confessing to the authorities is a strictly dominant strategy, making mutual defection the only equilibrium.

altruistic punishment, and sometimes even better. Cooperation, in turn, is strictly dominated by defection in the absence of altruistic punishment. This fact raises a question about evolutionary stability of cooperation and altruistic punishment in the presence of free-riders.

The model in Table 15 has interesting dynamics. The cooperative equilibrium gives both players higher utility than the mutual defection equilibrium, which makes it a salience equilibrium. At the same time, altruistic punishment is weakly dominated by cooperation (playing cooperation is always at least as good as playing altruistic punishment, and sometimes even better). Cooperation, in turn, is strictly dominated by defection in the absence of altruistic punishment. Later in the paper, I will analyze evolutionary stability of altruistic punishment using evolutionary game theory and the models of population dynamics.

Empirical examples of successful long-term management of the commons imply that some communities manage to ensure that the inequality $C > T - Y$ holds: following the rules of the community is more beneficial than free-riding and then being punished. Yet this may change. Field evidence also demonstrates that the cooperative equilibrium can de disrupted as a result of external factors such as natural cataclysms, infectious diseases, inter-group warfare, refugees, and state intervention. For example, droughts usually increase an individual incentive to free-ride and consume more water than others. The value of temptation is increased, and if more strict rules are not implemented the cooperative equilibrium is in danger (Ostrom 1990; Tang 1992; Mabry 1996).

The filed evidence suggests that under certain circumstances altruistic punishment can be a technical solution to the problem of free-riding or overexploitation. One of form of altruistic punishment is a widespread custom of "self-help" (Ellickson 1991), or self-enforcement of community rules by local means. In this case, individuals rely on personal retaliation as a primary countermeasure against deviants. As Ellickson put it, "a measured amount of self-help – an amount that would serve to even up accounts – is the predominant and ethically preferred response." (Ellickson 1991: 57). Although responsibility for the social control in many cases belongs to the victim, the individual has potential support from the whole group if the conflict escalates (Ostrom 1990; Anderson 1995). The cost of punishment generally increases with the frequency and degree of violation – producing sequentially warnings, gossip, damage of equipment, threat of violence, social ostracism, and even physical harm. Such customs do not lead to anarchy and vigilante justice. Actual cases of violence are quite rare, suggesting that the possibility of punishment is an effective deterrence mechanism (Ellickson 1991; Berkes et al 2001). Empirical examples suggest that successful commons have established sanctions against rule violations. In contrast, as Tang put it, "the communities without established self-sanctioning mechanisms have problems in rule conformance or maintenance" (Tang 1992).

The success of internal enforcement of cooperation is contingent upon a number of aspects. Below I examine four aspects, which appear to be critical for the prevention of the tragedy of the commons by means of altruistic punishment: (1) pre-commitment to punish free-riders, (2) possibility of cheating, or undetected defection, (3) open access regime as opposed to common-property regime, and (4) legal aspect of self-enforcement.

Pre-commitment

The model in Table 15 assumes that altruistic punishment is a stand-alone strategy, just like cooperation and defection[31]. Such strategy profile implies that the player choosing AP is pre-committed to punish defectors, despite the fact that such behavior will decrease his utility. Pre-commitment is critical for the cooperative equilibrium: if players *know* that one is pre-committed to punish defection, they will not defect. This makes pre-commitment beneficial in economic terms and adaptive in evolutionary terms. Therefore, the next question is: what are the effective pre-commitment mechanisms?

One class of such mechanisms are psychological adaptations such as emotions (Tooby and Cosmides 1992). One such emotion is anger (Fehr and Gachter 2002), which biologist Trivers called "moralistic aggression" (Trivers 1971). Although the act of punishment is irrational in economic sense, individuals know that emotional response may

---

[31] An alternative model is a two-stage game in which players choose cooperation or defection in the first stage and punishment or no punishment in the second (cf Axelrod 1986; Sethi and Somanathan 1996). In this case, the game is similar to the well-known chain-store paradox (Kreps and Wilson 1982).

override economic rationality. As a result, anger proves to be beneficial/adaptive for both players: interaction with a person who is capable of anger dramatically decreases the mutual temptation to free-ride.

Another class of effective mechanisms allowing pre-commitment are cultural adaptations. Communities around the world have moral codes, norms, and customs which encourage not only cooperation but also punishment of unfair behavior. In fact, such practice may take an extreme form: both cooperator and defector become the object of social ostracism if the cooperator fails to punish his offender[32]. In this case, pre-commitment is even unnecessary if non-punishing behavior is discouraged by the community in such an extreme way.

Finally, the issue of pre-commitment is solved by definition in the dynamic evolutionary model which I discuss below. In the model, cooperation, defection, and altruistic punishment are "types" which players do not choose in a traditional [optimizing] sense. Instead individuals act according to the type. If the utility, or fitness, of a type is greater than the population average then the proportion of this type increases in the next period. Such treatment of strategies in evolutionary game theory relaxes the assumption of hyper-rationality and optimizing behavior and solves the problem of pre-commitment.

Cheating

Cheating is a serious problem in the management of common-pool resources. If individual free-riding or over-use of the resources is difficult to observe, altruistic punishment is less likely to be effective. In terms of the model, the possibility of cheating is functionally the same as decreasing the cost of punishment *Y*, which increases individual incentive to free-ride. In the environment where defection is not observable, altruistic punishment becomes identical to cooperation, and I return to the standard Prisoners' Dilemma model.

---

[32] In terms of the model, it means that cooperators who face defectors are also punished for not being altruistic punishers.

Research in evolutionary psychology suggests that humans have evolved cheater-detection cognitive adaptations (Tooby and Cosmides 1992; Gigerenzer and Hug 1992; Orbell et al 2003). In other words, we are good at spotting free-riding. In particular, experiments on Wason selection task demonstrate that humans are capable of making complex logical inferences in a social environment where cheating is possible[33]. These psychological mechanisms, however, had evolved during the environment of evolutionary adaptedness (Pleistocene), or 99.9% of the human evolutionary time. As a result, such cognitive adaptations are generally effective for small groups of individuals (less than 100) who do not have sophisticated tools which can be used for cheating. In the environment, which is different from Pleistocene and which is characterized by big group sizes and new technologies, cheater-detection adaptations are less likely to be effective[34].

On the surface, such explanation of the problem of collective action is similar to Olson (1965) but the underlying reasons of the two arguments are different. According to the standard economic view, individuals in large groups are least likely to contribute because the benefit is diluted among members of the group. In the evolutionary framework, individuals are less likely to contribute because it is easier to get away with free-riding in a large group – a consequence of the nature of the environment in which human cognitive apparatus had evolved.

In response to the natural limits of cognitive cheater-detection, communities around the world have evolved sophisticated monitoring mechanisms designed to prevent cheating (Ostrom 1990; Bromley 1992; Tang 1992; Crean and Symes 1996; Alston, Libecap, and Mueller 1999). If cheating is still possible, I return to the traditional conclusion. To solve the problem of cheating in large groups, members of the commons need an external help such as new technologies or assistance from the state.

---

[33] Wason selection task also demonstrates that humans are not so good at solving abstract logical problems.

[34] Although new technologies make it easier to cheat, they also make it easier to monitor and detect the violation of rules, which makes it similar to an arms race.

Open access

Preceding discussion assumes that punishment of defectors is an appropriate response to defection and the question is whether or not such behavior is adaptive. In the real world, however, altruistic punishment is not always feasible. Successful self-enforcement of cooperative behavior is typically observed in the commons characterized by the common-property regimes, which are fundamentally different from open-access regimes (Ciriacy-Wantrup and Bishop 1975; Stevenson 1991; Ostrom 2001). Hardin's tragedy of the commons describes the tragedy of open access resources when there are no established norms, implicit or explicit community rules, no entry barriers for outsiders, and when monitoring and sanctioning of other members is problematic (Hardin 1968; Meltzer 1994; Berkes et al 1989; Kibreab 2002). In term of the theoretical model, it means that the punishment of defectors may be impossible or even inappropriate; therefore, management of open access resources is more challenging. As a consequence, one of the policy implications is that altruistic punishment can be a solution of the tragedy of open access resources only if the latter become common property.

Legal aspect: altruistic punishment and the state

Although under many circumstances state intervention proves to be necessary and beneficial, it may also have a surprisingly negative effect on commons management (Ostrom 2001; Smith 2003). Max Weber said in his famous lecture that the modern state seeks "to monopolize the legitimate use of force" (Weber 1958). Monopolization of force, however, undermines decentralized punishment by reducing its authority or even making it illegal. If the state alternatives to self-enforcement are not effective, the cooperative equilibrium is in danger. Field studies amply document how centralized intervention fails to preserve local customs and leads to the tragedy of the commons (Berkes et al 2001; Tang 1992). In Sudan as well as some other developing countries, the state rather than communal ownership has been the "major cause of inappropriate land use practices and consequently depletion of [common-pool resources]" (Kibreab 2002: 403). State intervention ironically may create more problems than benefits. As Vernon Smith (2003: fn. 52) put it:

> "… voluntary private associations for sharing the cost of a common good – policing – were subsequently undermined by statehood, and the publicly financed local sheriff as the recognized monopoly law enforcement officer. This observation contradicts the myth that a central function of government is to "solve" the free-rider problem in the private provision of public goods. Here I have the reverse: the incentive of the cattlemen's clubs was to free ride on the general taxpayer, assigning the sheriff the task of enforcing property rights in cattle."

Breaking the communal self-enforcement rules and norms could entail risk whereas re-creation of self-governing commons is a difficult task (Bromley 1992; Mabry 1996; Berkes et al 2001; Kibreab 2002).

Pre-commitment, successful monitoring, common-property regime, and a certain degree of independence are necessary attributes of the successful self-enforcement of cooperation in the commons. In the following section I examine the modified Prisoners' Dilemma as an evolutionary game. In this context, the problem of pre-commitment is solved by definition (strategies are types, not choices) and the problem of cheating is reflected in the parameters of the model[35]. Finally, if the commons are characterized by the open access regime and legal constraints on internal enforcement of cooperation, altruistic punishment is unlikely to be effective. An interesting link between the two factors was suggested by Elinor Ostrom. A large number of empirical cases show that "overexploitation of common-pool resources occurred when open access prevailed either because no set of individuals had property rights or because state property was treated as open-access property" (Ostrom 1992: 312).

Evolutionary stability of altruistic punishment

Take a closer look at the game in Figure 1. The fact that altruistic punishment can be a Nash equilibrium strategy does not necessary imply that it is also an evolutionary stable strategy, which is only a subset of Nash equilibria (Samuelson 1997). Before I proceed, it is important to define the concept of evolutionary stability. A strategy is evolutionary stable (ESS) when most members of the population adopt it and no single

---

[35] In terms of subjective expected utility, the cost of punishment for the defector becomes Xc, where c is the probability of getting away with defection.

individual has an incentive to adopt another strategy[36]. A "weaker" strategy is neutrally stable (NSS) when most members of the population adopt it but are indifferent between it and another strategy (Samuelson 1997).

If p is a proportion of the trait C in the population and q is the proportion of A, then:

(1.1) $U_C = p + q$

(1.2) $U_A = p + q - Y(1 - p - q)$

(1.3) $U_D = kp + (k - zY)q + m(1 - p - q)$

Defection is an evolutionary stable strategy against both cooperation and altruistic punishment: a rare cooperator or altruistic punisher will not survive in the population characterized by mutual defection ($p = 0$, $q \to 0$).

(2)
$$\begin{gathered} kp + (k - zY)q + m(1 - p - q) > p + q - Y(1 - p - q) \\ m > -Y \end{gathered}$$

The same is true, however, for altruistic punishment: in a community with a wide-spread norm of altruistic punishment, a rare defector will quickly disappear ($p = 0$, $q \to 1$).

(3)
$$\begin{gathered} kp + (k - zY)q + m(1 - p - q) < p + q - Y(1 - p - q) \\ k - zY < 1 \end{gathered}$$

Thus, both defection and altruistic punishment are mutually *evolutionary stable* against each other. Finally, in the absence of defection, cooperation and altruistic punishment have equal payoffs: each strategy is *neutrally stable* against the other ($p + q = 1$).

(4) $p + q = p + q - Y(1 - p - q)$

Notice that (non-punishing) cooperation weakly dominates altruistic punishment and at the same time it is strictly dominated by defection. Is the tragedy of the commons unavoidable? Weibull proves that in a 2*2 evolutionary game, a weakly dominated strategy

---

[36] This is a simplified definition. Formally, a strategy adopted by the most members of the population is called an *incumbent* strategy whereas other strategies are called *mutant* strategies.

will be extinguished from the population in the limit (Weibull 1996, Proposition 3.2). But what if the game is 3*3 and the weak dominance takes place only in the presence of a third strategy (defection)?

Although altruistic punishment is an evolutionary stable strategy against defection, it is indeed vulnerable to the *simultaneous* presence of defection and cooperation, even if the proportions of these strategies are small. The intuition is simple. In the presence of defectors, cooperation is more beneficial than altruistic punishment since cooperators do not incur the cost of punishment. As a result, the proportion of altruistic punishers relative to the proportion of cooperators will be decreasing over time. Eventually, there will be so few altruistic punishers and so many cooperators that a rare case of defection will be quickly imitated by other members of the commons as leading to a greater individual payoff. The analysis of replicator dynamics[37] of the Prisoners' Dilemma game confirms this intuition.

Formally, replicator dynamics can be described as follows. First, I need to find fitness (subjective expected utility) of each trait as function of the proportion of the traits in the population (1.1-1.3). Second, I find the average individual fitness in the population:

$$(5) \qquad \bar{U} = U_C p + U_A q + U_D (1 - p - q)$$

In classical model of replicator dynamics, the growth rates of the types are:

$$(6) \qquad \frac{\dot{p}}{p} = \left(U_C - \bar{U}\right), \qquad \frac{\dot{q}}{q} = \left(U_A - \bar{U}\right), \qquad \frac{\dot{w}}{w} = \left(U_D - \bar{U}\right),$$

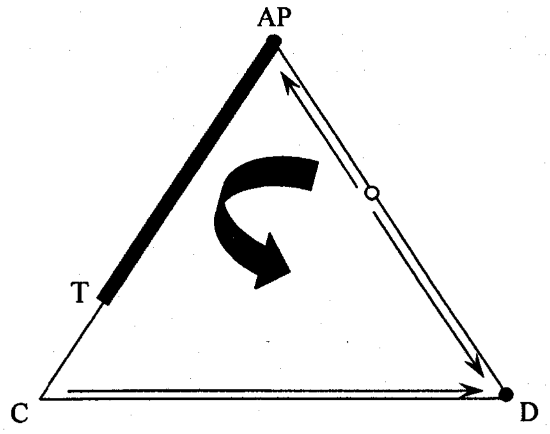
where $w = 1 - p - q$.

The simplex in Figure 15 is a convenient way to describe replicator dynamics of the strategies in the modified Prisoners' Dilemma game.

---

[37] Replicator dynamics can be though of a description of interaction between strategies over time. Formally replicator dynamics are a system of equations which represent the growth of individual traits in the population (Smith 1982; Samuelson 1997).

Figure 15. Replicator dynamics of the Prisoners' Dilemma game with altruistic behavior.



Note: The game has two equilibrium states: (1) D, when all members of the population defect, and (2) a basin of attraction AP-T, when the population is divided into cooperators and altruistic punishers. Notice that below the threshold T, the proportion of cooperators become so large relative to the proportion of altruistic punishers that a rare defection will lead to the tragedy of the commons.

Figure 15 demonstrates that in the short-run defection can be contained and the two surviving traits will be cooperation and altruistic punishment. The scenario is most likely when the following three conditions are satisfied:

1. Initial proportion of altruistic punishers is relatively high to the initial proportions of cooperators and defectors.
2. The parameters of the PD game do not discriminate against cooperators too much (e.g., the temptation to free- ride should not be much higher than the value of mutual cooperation).
3. The cost of punishment for the altruistic punisher should not be much higher than the cost of punishment for the defector.

Unfortunately for altruistic punishers and cooperators, the equilibrium state AP-T is not stable in the long run[38]. Every rare case of defection increases the proportion of cooperators and decreases the proportion of altruistic punishers. Eventually, the proportions of C and AP will move out of the equilibrium state and the population will unravel to mutual defection. Even if all members of the population are altruistic punishers, the defection strategy will eventually take over. A homogeneous population of altruistic punishers is immune to defection but it is not immune to the simultaneous invasion of cooperators and defectors. While defectors do not survive, cooperators do. In turn this changes the equilibrium state from all AP to the AP-T, which is not stable in the long run as I have established above.

What are the properties of convergence to the asymptotically stable state of mutual defection? A heterogeneous population of cooperators and altruistic punishers is invaded by defectors when individual fitness of free-riders is greater that the fitness of altruistic punishers. A greater initial proportion of altruistic punishers, q, in a population makes this scenario less likely. What is the minimal q that allows cooperation top survive? I know that the incumbent population consists of cooperators and altruistic punishers. I also know that free-riding is not adaptive as long as the fitness of defectors is not greater than the fitness of altruistic punishers. Hence, I have

$$(7)\qquad \begin{cases} kp+(k-zY)q+m(1-p-q)=p+q-Y(1-p-q) \\ p+q=1 \end{cases}$$

The system of equations has a unique solution:

$$(8)\qquad p=1-\frac{k-1}{zY},\qquad\qquad q=\frac{k-1}{zY}.$$

It is easy to confirm that the population growth of mutant defectors is zero at precisely the same value of q. $\frac{\dot{w}}{w}=\left(U_D-\bar{U}\right)=0$, when $q=\frac{k-1}{zY}$.

---

[38] Formally, equilibrium state AP-T is not *asymptotically stable* since rare defections (say, by mistakes or mutation) will eventfully move the proportion of the population below the threshold value *T* and the population will converge to the other equilibrium state – mutual defection. Mutual defection, on the other hand, is asymptotically stable. Rare cooperators or altruistic punishers will not be able to change it.

Since $k > 1$, the proportion of altruistic punishers must be $0 < q \leq 1$. Both greater z and $Y = \frac{X}{C}$ allow even for a relatively small proportion of altruistic punishers in the population to deter the invasion of free-riders. Whereas the effect of z on the likelihood of invasion is straightforward, the effects of X and C are almost counterintuitive. Greater cost of punishment for the punisher, X, *preserves* cooperation. Greater value of mutual cooperation, C, makes it *less likely* to survive.

I can also establish how big the weight on punishment has to be to deter free-riders. It clearly follows from the previous result (6) that the minimal weight should be $z = \frac{k-1}{qY}$. Greater k and C, and smaller q and X would require the weight on punishment to be very large in order to preserve cooperation. It is also interesting to determine the critical value of the temptation weight, k, that allows mutant defectors to outperform other types despite the punishment. Using the first equation in (5) I find that the critical value of the weight is $k = 1 + qzY$. Again, for the cooperation to survive, high value of k should be counter-balanced by a relatively high proportion of altruistic punishers, q, relatively large weight on punishment z, greater cost of punishment X, and smaller value of mutual cooperation C.

The observed replicator dynamics is ambiguous. A pessimist would notice that defection is the only surviving trait in the ultra-long run. An optimist would emphasize that the AP-T equilibrium state exists and is stable against defection except for the ultra-long run. The critical aspect of the dynamics is the slow gradual decrease of the proportion of altruistic punishers in the population. If the drift can be offset by some other factors, the commons will be stable against defection even in the ultra-long run. One such factor – absent in the current simple model – can be a conformist pressure to punish defectors (Henrich and Boyd 2001; Axelrod 1986). Another possibility is multi-level selection (Sober and Wilson 1998; Redondo 1993).

Evolutionary analysis of the Prisoners' Dilemma game with the altruistic punishment option has several advantages and is complementary to the standard approaches. Game theory tells us that the game has two Nash equilibria if $1 > k - zY$: (D,D) and (AP, AP). If $1 < k - zY$, there is only one Nash equilibrium: (D, D). Evolutionary game theory gives us the same predictions but provides more details. Mutual defection is always an asymptotically stable steady state whereas the altruistic punishment is not since it is vulnerable to the drift toward cooperation, which, in turn, is vulnerable to defection. Replicator dynamics gives us some intuition about the behavior of players out of equilibrium, or on the way to equilibrium. In the evolutionary model, I can also examine under what parameters the system will move from one steady state to another (when $q < \frac{k-1}{zY}$).

Altruistic punishment becomes very robust in the presence of group selection. The crux of the group selection argument is that a group of cooperators is more successful than a group of defectors in either implicit or explicit competition for resources. In this respect, each individual has incentive to cooperate to make his group stronger. On the other hand, each individual has also an incentive to defect to increase his payoff within a group. It is possible to show that in the model of group selection cooperation can survive if a number of conditions are met such as significant fitness differential between the groups, limited migration and genetic drift, extinction and/or formation of new groups. The evolution of cooperation under group selection is, nevertheless, very fragile. The main problem in such models is that within each group cooperation is *always* strictly dominated by defection for *all* possible parameters of the model. And this is where altruistic punishment differs from the [non-punishing] cooperation.

Altruistic punishment retains between-group advantage of cooperation but manages to avoid its within-group disadvantage. A homogeneous group of altruistic punishers is as successful as a homogeneous group of cooperators: members of both groups cooperate with each other. Both groups are more successful than a group characterized by defection. However, whereas the cooperative group is vulnerable to within-group defection, the group of altruistic punishers is not. As mentioned above, the success of altruistic punishment as a trait is directly proportional to its frequency. An increase in the proportion of altruistic punishers increases their own fitness and decreases the fitness of defectors. This is different from cooperation which is frequency independent within each group where it is always inferior to defection. The groups, in which cooperation is vulnerable to defection, will have a fitness advantage over the groups, in which cooperation is less vulnerable to defection. Vulnerability of defection, in turn, is a function of the proportion of altruistic punishers in the group. Hence, the groups with the higher proportions of altruistic punishers will have an advantage in the evolutionary competition for resources. Although this scenario can be supported by means of computer simulation models (cf Boyd et al 2003), future work should focus on the analytical characterization of the described dynamics.

It has been a folk theorem that punishment can sustain cooperation. In the traditional rational choice models, punishment meant defection in the context of repeated games. A possibility of continuing cooperation and punishment by some other means has been largely ignored as irrational in terms of subjective expected utility (but see the "Theory" column in Table 14). Moreover, costly punishment can be seen as a public good subject to the same problem of individual free-riding, which merely shifts the problem one step further. Nevertheless, experimental evidence and field studies suggest that altruistic punishment is a wide-spread phenomenon. Individuals are willing to incur the cost in order to punish defectors. In addition to experimental and field research, the phenomenon has a surprisingly large theoretical basis composed by the works from political science, economics, psychology, anthropology, computer science, and biology.

In my work, I made an attempt to explain why altruistic punishment should not be ignored. The overlap between the study of cooperation in social sciences and the research on evolutionary adaptedness of altruistic behavior in life sciences points out to the altruistic punishment as a possible technical solution to the survival of altruistic cooperation and to the tragedy of the commons. Further survey of the literature shows that scholars using different research *methods* (basic theory, experiments, and field studies) also provide evidence for the existence of the phenomenon. Using simple formal models from game theory and evolutionary game theory and evidence from evolutionary psychology, I made an attempt to integrated the interdisciplinary research and convince the reader that altruistic punishment is an important mechanism allowing for the escape from the Hobbesian world as well as altruistic behavior among non-human species. Costly self-enforcement of cooperation appears to be a part of human psychological apparatus as well as communal rules and norms. AP is rational in economic terms if precommitment is possible. It is also robust in evolutionary terms, especially in comparison with the non-punishing cooperation. Unlike cooperation, altruistic punishment is frequency dependent: the higher the proportion of altruistic punishers in the population the greater is their fitness. Although the trait is not asymptotically stable (i.e., unstable in the ultra-long run) within a single group, the problem can be offset by a conformist pressure within the group, or between-group competition. In the context of group selection, altruistic punishment retains the strength of cooperation in between-group competition and, at the same time, prevents defection from taking over within the group. In normative terms, altruistic punishment is a more desirable norm than the "traditional" *punishment by defection.* In the commons, Tit-for-Tat and other trigger strategies not only punish defectors but also harm individuals who cooperate, which may also lead to a new wave of defections. Altruistic punisher always cooperates and punishes only those who deserve it.

A possibility of altruistic punishment as a solution to the tragedy of the commons has immediate policy implications. If cheating can be prevented, if there is no emergency, and if the commons are characterized by the common-property regime, the community can be successful without external assistance such as state intervention. On the other hand, if monitoring is costly or impossible, if the commons are characterized by the open access regime[39], and if external factors such natural cataclysms are present, external assistance may be necessary. Failure to differentiate between the two cases as well as failure to recognize the importance of self-enforcement of cooperation as a community norm will only accelerate the opportunistic behavior and overexploitation of available resources. The tragedy of the commons often happens when individuals start treating their common-property resources as open access property. Recognizing altruistic punishment as a vital attribute of the common-property management may explain counter-productive policies.

---

[39] At the same time, the use of the term "commons" is questionable if the community is characterized by the open access regime as opposed to the common-property regime.

CHAPTER SEVEN

EVOLUTIONARY SIMULATION: MULTILEVEL SELECTION AND THE EVOLUTION OF HEROISM AS A FORM OF ALTRUISM

In the concluding substantive chapter of my work, I show how evolutionary simulation can allow us to analyze complex dynamic interactions both within and between groups. Archaeological (Keeley 1997; LeBlanc and Register 2003) and primatological (Wilson and Wrangham) evidence shows that our ancestral past involved persistent intergroup violence with significant mortality rates, raising the possibility of selection on altruistic propensities specialized for warfare. Here I report findings from one model allowing selection only on generalized altruism and another allowing selection on two forms of altruism: intra-group cooperation which, as conventional in models of generalized altruism, contributes to population growth; and "heroism," which contributes to success in intergroup violence. I show that there is only modest selection on "cooperation" but marked selection on "heroism" in a model that allows the two to evolve independently. "Heroism"—altruistic violence on behalf of one's group—may, therefore, be a fundamental and evolutionarily distinct form of human altruism.

An evolved willingness to fight for one's group despite the potential for serious fitness losses has been classically explained by the potential for even greater fitness gains to successful warriors (Tooby and Cosmides 1998; Eibl-Eibesfeldt 1979). Certainly, such fitness gains are well documented (Junker 1999; Chagnon 1988; Patton 2000; Ghiglieri 1999). Yet the potential that warfare, raids and opportunistic intergroup killing have for the survival of groups *as such* suggests that this willingness can also be addressed as a problem of multi-level selection.

Within the growing literature on group selection, the selective pressures from warfare as such have not been explicitly addressed, with willingness to fight for one's group being subsumed under "altruism"—the generalized tendency to contribute to the reproductive success of other group members at personal reproductive cost. Here I distinguish two forms of altruism: (1) "heroism," the privately costly act of fighting on the group's behalf; and (2) "cooperation," any other privately costly act of contributing to the reproductive success of group members. The task is to show whether the selective effects of intergroup violence operate differently on "heroism" and on "cooperation."

As Sober and Wilson have argued, a sufficient condition for the evolution of altruism via group selection is that "the differential fitness of groups (the force favoring altruists) must be strong enough to counter the differential fitness of individuals within groups (the force favoring the selfish types)." (Sober and Wilson 1998, p. 26.) For them, necessary conditions are: (1) there is a population of groups that vary in their proportion of altruistic types; (2) groups with altruists produce more offspring than groups without altruists; (3) although groups are isolated from each other, the progeny of both altruistic and non-altruistic groups must, at some stage, mix or otherwise compete in the formation of new groups. The model diverges from the latter requirements only insofar as groups compete explicitly via "warfare," and the populations of the respective groups form new groups based on the outcomes of such intergroup violence.

Intuitively, the model has the following general structure: A population of individuals is divided into several groups that occupy territories providing resources necessary for survival. In a succession of generations, individuals behave in ways that have fitness consequences for themselves and others. At the end of a generation, individuals' reproductive success depends on their accumulated fitness, and offspring resemble their parent within a specified range of variation. During their life, individuals can increase the reproductive success of their group via altruistic cooperation (the greater the incidence of cooperation, the faster the growth of the group) which comes at a personal reproductive cost.

War occurs when one group's population exceeds the natural resource capacity of its territory; the group then attacks another group, randomly chosen. The outcome is determined by the number of individuals who participate in the fighting and the degree to which each participant behaves "heroically," interpreted as exposing oneself to danger on behalf of one's group. Fighting heroically is an altruistic act, increasing the group's potential reproductive success at an individual reproductive cost (the hero's increased probability of dying without offspring). A smaller group can win a war against a larger group if its members are more heroic, but—ceteris paribus—larger groups have an advantage. Members of the losing group are annihilated, and the winning group fissions, with some members moving to occupy the newly vacated territory. In the next generation, the newly formed group is capable of either initiating hostilities or responding to attack.

I have two particular models. The first treats "altruism" as a generalized propensity; altruists cooperate during peace and participate in intergroup violence. The second separates the propensity for intragroup "cooperation" from the propensity for intergroup "heroism." Cooperation contributes to the group's general reproductive success while heroism contributes directly to the group's fighting strength in war. I compare outcomes of the two models to separate the effects of selective pressures on cooperation and heroism.

Formally: The world in the simulation consists of several territories with fixed boundaries. A territory $g = 1, 2, \ldots$ is endowed with R_g natural resources. A number of individuals who populate this territory, n_g, form a group consuming its resources. Survival of the group's members is secured if resources per capita do not fall below an exogenous survival threshold, $T : R_g / n_g > T$; otherwise, the group attacks another group chosen at random.

Members of all groups have the same baseline fertility, $P_b \in [0,1]$, a natural probability of having offspring in a generation. Baseline fertility can be adjusted for each group depending on its members' average cooperation. With the latter incorporated, the resulting baseline for the *group* g becomes $P_b + P_c A_g$, where $P_c > 0$ is the difference between purely cooperative and non-cooperative groups, and $A_g = \frac{1}{n}\sum_{i=1}^{n_g} c_i$ is the group's average cooperation with $c_i \in [0,1]$ being the propensity to cooperate of a member i. Individual cooperation is costly to fertility, decreasing the probability of having an offspring by $x > 0$; the cost is assessed proportional to c_i. If a group's resources per capita fall below the survival threshold, it attacks another group chosen at random. Groups must fight if attacked. Each member of a group can increase that group's fighting strength by active fighting in a war—which I call heroism. Similar to cooperation, I assume that heroism decreases the probability of having an offspring by $y > 0$. Heroism is also a continuous variable, $h_i \in [0,1]$. The group with the greater fighting strength, $\sum_{i=1}^{n_g} h_i$, wins. Members of the losing group die, or otherwise vacate the territory; the winning group fissions, dividing its members between its original territory and the vacated territory so that resources per capita are equal in the two territories. Combining the factors above, the probability of member *i* of a *surviving* group having an offspring is: $P_b + P_c A_g - c_i x - h_i y w$, where $w = 0,1,...$ is the number of wars. Thus, *within* a single group more cooperative individuals always have a lower probability of reproducing. The same is true for more heroic individuals if $w \neq 0$. However, groups whose members are more cooperative and heroic grow faster and are more likely to win wars. The one attribute model is identical except that $c_i \in [0,1]$ is *both* a propensity to cooperate *and* to participate in a war.

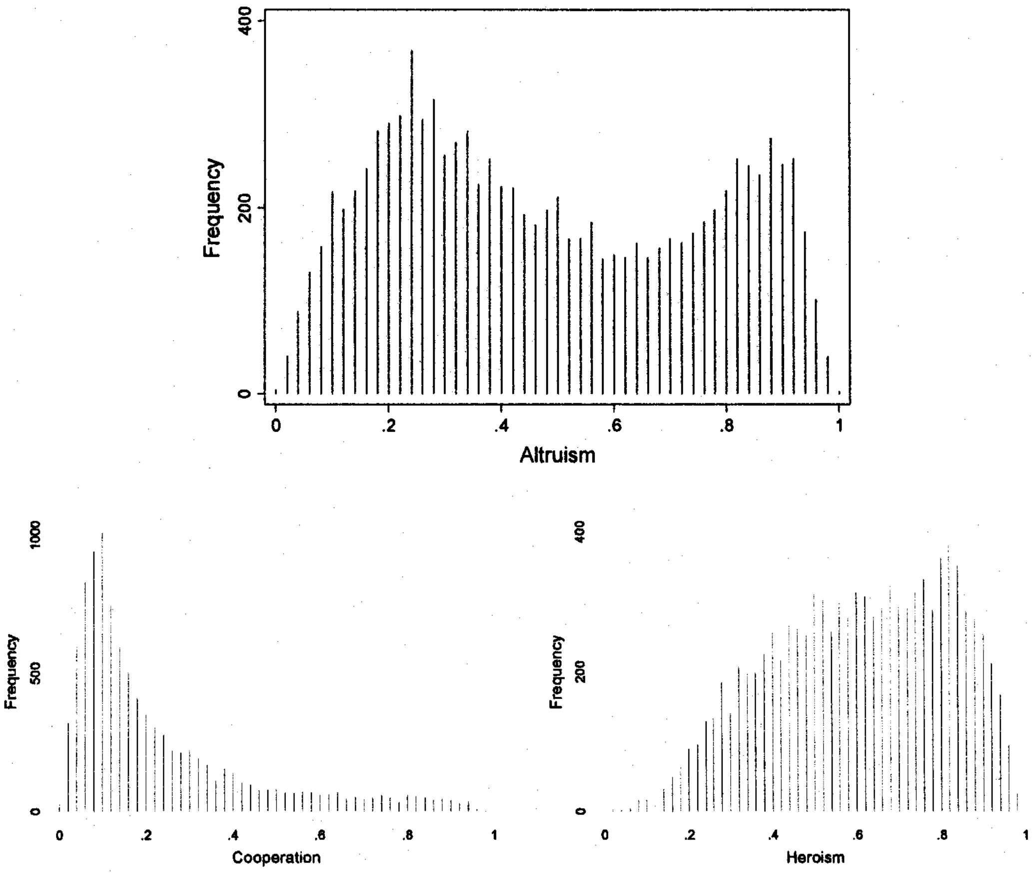
In each generation an individual i may reproduce depending on its probability of having an offspring. Reproduction is asexual and stochastic. Offspring attributes are copied from the parent subject to a uniform shock with zero mean and an exogenous variance. In addition, with a very small exogenous probability, offspring attributes can be drawn from a uniform distribution [0,1].

To examine selection on altruism, cooperation, and heroism I conducted 10,000 simulations for both versions of the model. Parameters for each run were drawn from a multivariate uniform distribution of feasible values with an identity correlation matrix such that $P_b \sim [0,0.5]$, $P_c \sim [0,0.5]$, $x \sim [0,0.15]$, $y \sim [0,0.15]$, $R_g \sim [50,100]$ (with the starting population being $n_g = R_g / 2$). Several parameters were exogenously fixed: $g = 4$, $T = 1$, mean starting $c_i = 0.5$ and $h_i = 0.5$; this did not affect the substantive results. For each run, I recorded simulation parameters and an equilibrium level of the dependent variables (altruism, cooperation, heroism, number of wars) approximated by the respective moving averages for the last 1,000 generations of 2000 in each simulation. Increasing the number of generations for each run does not affect the substantive results.

Selection on altruism (model 1) and on cooperation and heroism (model 2) occurs in the environment where individuals interact, and this is defined by the following parameters: Number of groups; amount of natural resources in the groups' territories; baseline individual fertility; extent to which cooperation can improve fertility; reproductive costs to individuals and initial distribution of cooperation and of heroism; mutation rate; and the number of generations. I ran 10,000 simulations of each of the two models. Parameters for each run were drawn from a random distribution of feasible values allowing us to include a wide variety of possible environments in the statistical analysis. To approximate the equilibrium values of the dependent variables, I calculated the mean values of altruism (model 1) and of cooperation and heroism (model 2) for the last 1,000 of 2000 generations in each simulation.

Figure 1 shows the frequency distribution of values of altruism, cooperation, and heroism for the last 1000 generations. Altruism has a bimodal distribution (mean 0.48), implying that the parameter space can be partitioned into two types of environments—one with strong selection pressure favoring altruism and one without it.

Figure 17. Frequency distributions of altruism, cooperation, and heroism.



The spike plots are frequency distributions of equilibrium values of, respectively, altruism, cooperation, and heroism for 10,000 runs of the simulation. Environmental parameters for each run were drawn from a random distribution of feasible values (see Methods). In both versions of the model, altruism reduces individual fertility compared to others *within* a group. Groups with a higher incidence of altruism, however, grow faster and are more likely to win wars.

Separating cooperation and heroism as independent components of altruism leads to a very different pattern. Cooperation has a mean of 0.23 and a median of 0.15; for a wide range of parameters, there was very little selection on that attribute. Heroism, on the other hand, has a mean of 0.61 and a median of 0.62; for most of the parameter space explored, there was strong selection on this component of altruism. Since there was no correlation between cooperation and heroism (r = 0.011), the bimodal distribution of altruism in the first model appears to represent two types of environments—those with and without strong selective pressure on heroism. Cooperation meanwhile has only a marginal impact on the evolution of altruism. Statistical analysis of the multiple simulation runs (Table 16) provides further evidence that cooperation and heroism are two distinct forms of altruism, subject to distinct adaptive pressures. *Growth difference* describes the growth advantage that more cooperative groups have over less cooperative groups. In general, the larger the growth difference, the sooner more cooperative groups will become overpopulated, experience resource stress, and attack other, possibly less populated, groups. In the two attribute model this variable has a significant positive effect on the evolution of cooperation, but a significant negative effect on the evolution of heroism. In the one attribute model, however, growth difference appears to have no effect on the evolution of altruism. The reason for this difference is that, in the two-attribute case, wars can be decided by a difference in group size that swamps the impact of heroism—viz., a large group of not-so-heroic fighters can defeat a small band of more heroic ones. In the one attribute case, however, generalized altruism captures both (opposing) effects, making the growth difference falsely appear insignificant.

The *number of wars* has a strongly positive effect on the selection of altruism, which turns out to be a consequence of selection on cooperation, not heroism. Although warfare *per se* is critical to the evolution of heroism, the *number* of wars does not affect the equilibrium level of this attribute; a relatively small number of wars selects for heroism as strongly as a large number of wars. But cooperation is quite elastic to the number of wars; a 1% change in war frequency increases cooperation by 0.4%. Thus heroism can evolve even with only intermittent warfare while cooperation is more likely to evolve in an environment of persistent intergroup violence.

Table 16. Effect of parameters of environment on altruism, cooperation, and heroism.

Dependent Variable	Altruism	Cooperation	Heroism
Growth Difference	0.003	0.061**	-0.008*
Wars	0.497**	0.404**	-0.001
Cost (Altruism)	-0.373**		
Cost (Cooperation)		-0.572**	-0.001
Cost (Heroism)		-0.003	-0.183**
Resources Mean	-0.322**	-0.513**	-0.115**
Resources Std. Dev.	-0.170**	-0.049**	-0.144**
Constant	-0.381	-1.135**	-0.265
Observations	9950	9919	9919
R-squared	0.63	0.55	0.19

Dependent and independent variables are analyzed in the form of natural logarithms. In double log form, regression coefficients can be interpreted as elasticities measuring the percentage change in the dependent variable for a one percent change in an independent variable with other variables constant. Other statistical specifications lead to substantively similar results. One star indicates significance at 5%; two stars indicate significance at 1%. Cases when the dependent variable was zero were dropped from the double log specification.

Not surprisingly the size of the cost of any altruistic behavior has a negative effect on the altruistic trait. Interestingly, cooperation appears to be three times more elastic to the cost (-0.572) than heroism (-0.183). Altruism in the one attribute models appears to capture a middle ground between the two attributes and has elasticity (-0.373). The fact the cooperation is much more elastic to the cost than heroism is another evidence that the two attributes differ in their sensitivity to the environment. The result, further, suggests that heroism is a more robust form of altruistic behavior since it is relatively inelastic to the cost. Cooperative behavior, on the other hand, is rather vulnerable to the size of the cost.

Cooperation is also several times more elastic to the *mean* of groups' resources (-0.513) than is heroism (-0.115). A group's resources determine how large it can grow, and in a larger group, one individual necessarily has a smaller impact on a group's mean level of cooperation. The marginal impact of an individual's heroism is also smaller in a larger group, but heroism's marginal impact is more critical when groups are similarly sized. Indeed, heroism is almost three times more elastic to the *standard deviation* of groups' resources than is cooperation. When the standard deviation of groups' resources is large (hence groups vary more in size), wars are decided by numbers instead of by heroism.

In short, cooperation is more likely to evolve if groups are small, whereas heroism is more likely to evolve if they are similar in size. Analysis of the *distribution* of resources in an ecology inhabited by multiple groups is, therefore, also consistent with cooperation and heroism being evolutionarily distinct forms of altruism.

Nothing in the present analysis challenges standard ideas explaining heroism in terms fitness advantages to heroic individuals; as indicated earlier, the anthropological data supporting those ideas are strong. But the findings show that heroic dispositions could also have evolved to significant levels as a result of warfare-based between-group selection, and have evolved by that means on a trajectory distinct from the one followed by other forms of altruism.

The formal evolutionary model presented in this chapter is an example how we can analyze complex dynamic environment with several endogenous variables (altruism, cooperation, heroism, number of wars, population size) and large number of environmental parameters. The complexity is too overwhelming for an analytical approach. Although the computational model allows us to deal with the complexity, some generality is inevitably lost. To alleviate the problem (somewhat), one can use multiple simulation runs for a wide range of randomly drawn parameters. Generated data then can be used for a more systematic statistical analysis of the properties of the environment and their influence on the variables of interest.

Political scientists are often disconcerted by the "too much oversimplification" in the formal political theory. Along with the chapter on elections and turnout, this chapter shows how we manage complexity without losing some of the rigor of logic and mathematics. Highly complex models of political science can be created and analyzed if the underlying interactions are presented by means of the simple evolutionary microfoundations such as adaptive learning, imitation, and replicator dynamics.

CHAPTER EIGHT
CONCLUSION

The primary goal of my work has been to show that formal evolutionary modeling can be applied in political science as a useful, perhaps complementary, methodology. In the conclusion I will summarize the key attributes and benefits of the respective formal evolutionary models and how they can be applied in political science. But before that I would like to point out that, although my focus has been primarily methodological, the project also has yielded a number of substantive results.

First, I show that voting behavior in small legislative bodies cannot be fully captured by the orthodox tools of rational choice theory. Experimental evidence suggests that individuals will make rational choices but they may not necessarily have rational expectations about the behavior of others. Such behavior and beliefs have dramatic implications for the way we model legislative politics. Strategic voters may not believe that others are strategic. As a result the choice of agendas – the rules of the game – can be very different from the one predicted by the orthodox rational choice theory. Evolutionary models can describe the dynamics of the process as well as the convergence to the Nash equilibrium in the limit.

Second, I show how politicians offer divergent platforms and why voters choose to go to the polling places despite being unlikely to be pivotal. The models of adaptive learning and evolutionary imitation allow us to avoid some strong assumptions (e.g., that a voter knows everybody else's behavior in the country). In addition, these models produce empirically realistic results as well as novel insights (see Table 13 for the summary). Some of the results turned out to be completely unexpected. For example, it is the first model to suggest that ideological segregation reduces turnout and yields polarized parties.

Third, I examine altruistic punishment as a potential solution to the tragedy of the commons. The phenomenon has become very popular recently across social and life sciences. Using evolutionary game theory I show altruistic punishment is robust in the short run but asymptotically unstable in the long run. For altruistic punishment to be a viable long-term solution, there must be some additional benefit to the bearer of the trait or to the commons where altruistic punishment is prevalent. On the individual level, such benefit may be due to conformist pressure (i.e., not having to pay the pressure cost), or due to punishment of non-punishers and no play option (Fowler 2005). However, we may not have to make those additional assumptions if we allow for explicit or implicit competition among groups.

Intergroup competition underlies any multilevel selection model. This is the fourth substantive area that I examine here. Substantive insights in the respective chapter include the fact that strictly altruistic (cost to oneself, benefit to others) can evolve due to multilevel selection. Moreover, we can distinguish different forms of altruistic behavior such as cooperation and heroism. The evolutionary simulation model suggests that evolutionary trajectories of cooperation and heroism are very different despite the fact that mathematically the two fit the definition of altruistic behavior. Heroism as a trait appears to be much more robust than cooperation in the environment of frequent intergroup violence. Finally this chapter provides predictions how exogenous environmental variables such as the level of available resource and resources standard deviation may affect the endogenous variables: cooperation and heroism.

The insights discussed above are arguably important but very different given their substantive area: agenda-setting, turnout, platforms, altruistic punishment, intergroup conflict. What unites them, however, is the formal evolutionary methodology. This is the main focus of my dissertation. The fact that the substantive areas are so different should be viewed as a strength and relevance of the evolutionary approach in the discipline of political science.

All evolutionary models can be divided into two categories: (1) when evolution is used as a an analogy/metaphor for modeling purposes, (2) when evolution is used as a homology, i.e., having implications for the actual cognitive design of human brain, the "wet structure." The models of agenda-setting and elections obviously fit the first category.

The models of altruistic punishment and intergroup violence may fit the second. The main criterion is whether or not the environment in the model approximates the environment of evolutionary adaptedness – humans' ancestral past. The first category of models may relax certain assumptions, yield novel results, and describe the out-of-equilibrium dynamics, but the emphasis will be on the observed behavior and outcomes, and *not* the cognitive design of the brain of the participating individuals. The second category of models, however, may have important implications about innate propensities for certain types of behavior as an actual product of biological evolution.

Table 17. Dissertation summary table (substantive chapters).

Substantive area	Level of Analysis	Evolutionary Tools	Methodological Contribution
Endogenous agenda-setting: strategic voting over alternatives and a corresponding choice of agendas	Analogy	Adaptive learning Evolutionary imitation	Relaxing the assumption of hyperrationality Analysis of out-of-equilibrium dynamics Consistency with Nash equilibrium in the limit
Model of elections: political competition and voters' turnout	Analogy	Adaptive learning Evolutionary imitation	Relaxing the assumption of hyperrationality Modeling elections as a dynamic process
Altruistic punishment as a behavioral type	Homology	Evolutionary game theory	Analysis of evolutionary stability and convergence trajectories
Cooperation and heroism as domain-specific forms of altruism	Homology	Evolutionary computation	Analysis of intergroup and intragroup interactions in a complex environment

Table 17 summarizes the four substantive areas that I considered here. In my work, I suggest that formal evolutionary modeling applied in political science has several important benefits.

First of all, evolutionary models allow us to relax the assumption of individual hyper-rationality and unlimited information-processing power. Formal political theory is often criticized that the underlying model of a man is unrealistic, and, as a result, all the consequent analysis if fundamentally flawed. Evolutionary models solve the problem in two ways. In case of adaptive learning and evolutionary imitation, individuals use very simple heuristics and informational shortcuts to make their decisions. In case of evolutionary game theory and evolutionary computation, strictly speaking individuals do not make any choices altogether: instead they are born as types and cannot change their behavior. What does change is the composition of the population such that the most successful types are most likely to replicate.

Relaxing the assumption of hyper-rationality may be seen as a major departure from the classical game theory. In reality it is not. Formal evolutionary models allow us to have the cake and eat it too: the models rely on logic and mathematics and in the limit lead to the Nash equilibrium solution (but not all Nash equilibria can be evolutionary outcomes). Most importantly, the evolutionary approach provides insights into out-of-equilibrium behavior and how a particular steady state can be achieved if such states exist. The latter attribute of formal evolutionary models should be especially useful for political scientists. The social and political environment changes all the time, as a result many individual decisions are incremental adjustments – as is the evolutionary process. The tools of adaptive learning and evolutionary imitation allow us to model incremental adjustment and social learning (as exemplified in the chapter devoted to elections and turnout). This becomes particularly important if the environment and interactions in the model are highly complex (as exemplified in the chapter on multi-level selection). Evolutionary models allow us to analyze complex social processes on the basis of very simple local interactions and very simple defendable logic.

Finally, formal evolutionary models may provide insights into the actual cognitive structure of the brain (cf Orbell et al 2005). If the environment in the model is closely related to the environment of evolutionary adaptedness, one may suggest that predictions of the model may correspond to the actual biological evolution. For example, if our ancestors lived in small groups competing with each other for resources then it is possible that the altruistic propensity to defend one's group called "heroism" can be an innate characteristic of humans.

My task was to show that formal evolutionary modeling can help political scientists in general – and formal political theorists in particular – to create better and more interesting models which yield empirically realistic predictions about highly complex social and political behavior. By applying formal evolutionary models in four different substantive areas within political science I managed to produce novel substantive results, which would be difficult (if possible at all) to obtain by means of other models and/or approaches.

BIBLIOGRAPHY

Aldrich, John H. 1993. "Rational Choice and Turnout." *American Journal of Political Science* 37:246-278.

Alexander, R. D. *The Biology of Moral Systems* (Aldine de Gruyter, New York, 1987).

Alston, Lee J., Gary D. Libecap, and Bernardo Mueller. 1999. *Titles, Conflict, and Land Use.* Ann Arbor: The University of Michigan Press.

Anderson, Terry L., and Fred S. McChesney. 2003. *Property Rights: Cooperation, Conflict, and Law*. Princeton: Princeton University Press.

Anderson, Terry L., and Randy T. Simmons, eds. 1993. *The Political Economy of Customs and Culture*. Lanham: Rowman and Littlefield Publishers.

Anderson, Terry. L. *Sovereign Nations or Reservations?* (Pacific Research Institute for Public Policy, San Francisco, 1995).

Arbuckle, M., and M. Metzger. 2000. *Food for Thought: A Brief History of The Future of Fisheries Management*. Challenger Scallop Enhancement Company Limited.

Arrow, Kenneth. 1963. *Social Choice and Individual Values.* 2nd ed. New Haven: Yale University Press.

Austen-Smith, David. 1987. "Sophisticated Sincerity: Voting Over Endogenous Agendas." *American Political Science Review*, 81:1323-1330.

Axelrod, R. & Hamilton, W.D. 1981. The evolution of cooperation. *Science* 211, 1390-1396.

Axelrod, R. 1986. "An Evolutionary Approach to Norms." *Am. Pol. Sci. Rev.*, 80:1095-1111.

Axelrod, Robert. 1984. *The Evolution of Cooperation.* Basic Books.

Axelrod, Robert. 1997. *The Complexity of Cooperation: Agent-Based Models of Competition and Collaboration*. Princeton, NJ: Princeton University Press, 1997.

Banks, Jeffrey S. 1985. "Sophisticated Voting Outcomes and Agenda Control." *Social Choice and Welfare*, 1:295-306.

Baron, David P., and John Ferejohn. 1987. "Bargaining and Agenda Formation in Legislatures." *American Economic Review*, 77:303-309.

Bartlett, Randall. *Economic Foundations of Political Power*. New York: The Free Press, 1973.

Bateson, Melissa. 2002. "Recent Advances in Our Understanding of Risk-Sensitive Foraging Preferences." *Proceedings of the Nutrition Society*. 61:509-516.

Beck, Paul Allen, Russell J. Dalton, Steven Greene, and Robert Huckfeldt. 2002. "The Social Calculus of Voting: Interpersonal, Media, and Organizational Influences on Presidential Choices." *American Political Science Review* 96:57-74.

Becker, Gary S. *The Economic Approach to Human Behavior*. Chicago: The University of Chicago Press, 1976.

Berch, Neil. 1993. "Another Look At Closeness and Turnout - The Case of The 1979 and 1980 Canadian National Elections." *Political Research Quarterly* 46:421-432.

Berelson, Bernard, Paul F. Lazarsfeld, And William N. McPhee. 1954. *Voting*. Chicago: University of Chicago Press.

Berkes, F., Feeny, D., McCay, B. J., & Acheson, J. M. The benefits of the commons. *Nature* 340, 91-93 (1989).

Berkes, F., Mahon, R., McConney, P., Pollnac, R. & Pomeroy, R. *Managing Small-scale Fisheries* (International Development research Centre, Ottawa, 2001).

Binmore, Ken. 1992. *Game Theory and Social Contract. Volume 2: Just Playing*. Lexington: D.C. Heath.

Black, Duncan. 1958. *Theory of Committees and Elections*. Cambridge: Cambridge University Press.

Boyd Robert, Herbert Gintis, Samuel Bowles, and Peter J. Richerson. 2003. "The Evolution of Altruistic Punishment," *Proceedings of the National Academy of Sciences*, v. 100, 6:3531-3535.

Boyd, R. & Richerson, P. J. *The Origin and Evolution of Cultures* (ed. Sitch, S.) (Oxford University Press, Oxford, 2005).

Boyd, R. & Richerson, P.J. 1992. Punishment allows the evolution of cooperation (or anything else) in sizeable groups. *Ethol. Socibiol.* 13, 171-195.

Boyd, Robert, and Peter J. Richerson. 1985. *Culture and the Evolutionary Process.* Chicago: University of Chicago Press.

Brody, Richard A. and Benjamin I. Page. 1973. "Indifference, Alienation and Rational Decisions: The Effects of Candidate Evaluations on Turnout and the Vote." *Public Choice* 15:1-18.

Bromley, D. W. (ed) *Making the Commons Work* (Institute for Contemporary Studies, San Francisco, 1992).

Buchanan, James M., and Gordon Tullock. *The Calculus of Consent.* Indianapolis: Liberty Fund, 1962 (1990 ed.).

Busch, Marc and Eric Reinhardt. 2000. "Geography, International Trade, and Political Mobilization in US Industries." *American Journal of Political Science* 44:4, 703-719.

Calvert, Randall L., and Richard F. Fenno, Jr. 1994. "Strategy and Sophisticated Voting in the Senate." *The Journal of Politics*, 56:349-376.

Calvert, Randall. 1985. "Robustness of the Multidimensional Voting Model, Party Motivations, Uncertainty and Convergence," *American Journal of Political Science*, 39: 69-95.

Campbell, Angus, Gerald Gurin, and Warren E. Miller. 1954. *The Voter Decides.* Evanston: Row, Peterson and Company.

Cassel, Carol A. 1999. "Voluntary Associations, Churches, and Social Participation Theories of Turnout." *Social Science Quarterly* 80:3, 504-517.

Chagnon, N. Life Histories, Blood Revenge, and Warfare in a Tribal Population. *Science* 239, 985-992 (1988).

Ciriacy-Wantrup, Siegfried V., and Richard C. Bishop. 1975. "Common Property as a Concept in Natural Resource Policy." *Natural Resources Journal* 15:713-727.

Clutton-Brock, T. H., and G. A. Parker. 1995. "Punishment in Animal Societies." *Nature* 373:209-216.

Cohen, Linda, Michael E. Levine, and Charles R. Plott. 1978. "Communication and Agenda Influence: The Chocolate Pizza Design," in Heinz Sauerman, ed., *Coalition Forming Behavior, Contributions to Experimental Economics*, 8:329-57. Tubingen, Germany: Mohr.

Conley, Patricia. 2001. *Presidential Mandates: How Elections Shape the National Agenda.* Chicago: University of Chicago Press.

Cosmides, Leda, and John Tooby. "Better than Rational: Evolutionary Psychology and the Invisible Hand." AEA Papers and Proceedings. May 1994, pp. 327 – 332.

Cosmides, Leda, and John Tooby. 1996. "Are Humans Good Intuitive Statisticians After All? Rethinking Some Conclusions from the Literature on Judgment Under Uncertainty." *Cognition.* 58, 1 – 73.

Cox, Gary W. and Munger Michael C. 1989. "Closeness, Expenditures, and Turnout in The 1982 United-States House Elections." *American Political Science Review* 83:1, 217-231.

Cox, Gary. 1984. "An Expected Utility Model of Electoral Competition." *Quality and Quantity* 18, 337-349.

Crean, Kevin, and David Symes (eds.) 1996. *Fisheries Management In Crisis.* Oxford: Fishing News Books.

Davis, Otto, M. J. Hinich and P. C. Ordeshook, 1970. "An Expository Development of a Mathematical Model of the Electoral Process." *American Political Science Review* 64:426-448.

Denzau, Arthur T., and Robert J. Mackay. 1983. "Gatekeeping and Monopoly Power of Committees: An Analysis of Sincere and Sophisticated Behavior." *American Journal of Political Science*, 27:740-61.

Downs, Anthony. 1957. *An Economic Theory of Democracy.* Boston: Addison Wesley.

Eckel, Catherine, and Charles A. Holt. 1989. "Strategic Voting in Agenda-Controlled Committee Experiments." *American Economic Review*, 79:763-773.

Eibl-Eibesfeldt, I. *Biology of Peace & War* (Viking Press, New York, 1979).

Ellickson, Robert C. 1991. *Order without Law: How Neighbors Settle Disputes.* Cambridge: Harvard University Press.

Elster, Jon. 1989. *The Cement of Society: A Study of Social Order.* Cambridge: Cambridge University Press.

Elster, Jon. *Sour Grapes: Studies in the Subversion of Rationality.* Cambridge: Cambridge University Press, 1983.

Elster, Jon. *Ulysses and the Sirens: Studies in Rationality and Irrationality*. Cambridge: Cambridge University Press, 1979.

Enelow, James M. 1981. "Saving Amendments, Killer Amendments, and an Expected Utility Theory of Sophisticated Voting." *The Journal of Politics*, 43:1062-89.

Farquharson, Robin. 1969. *Theory of Voting*. New Haven: Yale University Press.

Fehr, Ernst, and Rockenbach, B. "Detrimental effects of sanctions on human altruism." *Nature*, 422, 137 – 140.

Fehr, Ernst, and Simon Gachter. 2002. "Altruistic punishment in humans." *Nature*, 415:137-140.

Fenster, Mark J. 1994. "The Impact of Allowing Day of Registration Voting On Turnout in United-States Elections From 1960 To 1992 - A Research Note." *American Politics Quarterly* 22:1, 74-87.

Ferejohn, John and Morris Fiorina. 1974. "The Paradox of Not Voting: A Decision Theoretic Analysis," *American Political Science Review* 68: 525-536.

Ferejohn, John. *Pork Barrel Politics.* Stanford: Stanford University Press, 1974.

Field, Alexander J. 2001. *Altruistically Inclined? The Behavioral Sciences, Evolutionary Theory, and the Origins of Reciprocity.* Ann Arbor: The University of Michigan Press.

Fiorina, Morris, and Charles R. Plott. 1978. Committee Decisions Under Majority Rule: An Experimental Study. *American Political Science Review*, 72:575-598.

Fishburn, Peter C. "Paradoxes of Voting." *American Political Science Review*. 68 (June 1974), pp. 537 – 546.

Fiske, John. "Cultural Studies and the Culture of Everyday Life," in L. Grossberg, C. Nelson, P.A. Treichler, *Cultural Studies*. New York: Routledge, 1992.

Fox, Robin. "Prejudice and the Unfinished Mind: A New Look at an Old Failing." *Psychological Inquiry*. 3(2) 1992, pp. 137 – 152.

Frank, Steven A. 1995. "Mutual Policing and Repression of Competition in the Evolution of Cooperative Groups." *Nature* 377:520-522.

Friedman, Daniel. 1991. "Evolutionary Games in Economics." *Econometrica*, 59:637-666.

Friedman, Daniel. 1996. "Equilibrium in Evolutionary Games: Some Experimental Results." *The Economic Journal*, 106:1-25.

Friedman, Daniel. 1998. "On Economic Applications of Evolutionary Game Theory." *Evolutionary Economics*, 8:15-43.

Fudenberg, Drew, and Eric Maskin. 1986. "The Folk Theorem in Repeated Games with Discounting and Incomplete Information." *Econometrica* 54:533-554.

Gaulin, Steven J.C., and Donald H. McBurney. *Psychology: An Evolutionary Approach.* Prentice-Hall, 2001

Ghiglieri, M. P. *The Dark Side of Man: Tracing the Origins of Violence* (Perseus, Reading, Mass, 1999).

Gigerenzer, Gerd, and Hoffrage, U. "How to Improve Bayesian Reasoning Without Instruction." *Psychological Review.* 102 (1995), 684 – 704.

Gigerenzer, Gerd, and Hug, K. 1992. "Domain Specific Reasoning: Social Contracts, Cheating, and Perspective Change." *Cognition*, n. 43, 127 – 171.

Gigerenzer, Gerd, et al. 1999. *Simple Heuristics that Make Us Smart.* Oxford University Press.

Gigerenzer, Gerd, U. Hoffrage, and H. Kleinbolting. "Probabilistic Mental Models: A Brunswikian Theory of Confidence." *Psychological Review.* 98 (1991), pp. 506 – 528.

Gigerenzer, Gerd. "How to Make Cognitive Illusions Disappear: Beyond "Heuristics and Biases." In W. Stroebe and M. Newstone (Eds.), *European Review of Social Psychology.* Chester: Wiley, 1991, pp. 83 – 115.

Gigerenzer, Gerd. "On Narrow Norms and Vague Heuristics: A Reply to Kahneman and Tversky." *Psychological Review.* 103 (1996), pp. 592 – 596.

Gigerenzer, Gerd. "The Bounded Rationality of Probabilistic Mental Models." In K.I. Manktelow and D.E. Over (Eds.), *Rationality.* London: Routledge, 1993, pp. 284 – 313.

Gigerenzer, Gerd. "Why the Distinction Between Single-Event Probabilities and Frequencies is Relevant for Psychology and Vice Versa." In G. Wright and P. Ayton (Eds.) *Subjective Probability.* New York: Wiley, 1994, pp. 129 – 162.

Gigerenzer, Gerd., and Peter M. Todd. "Fast and Frugal Heuristics: The Adaptive Toolbox" in Gerd Gigerenzer, et al. *Simple Heuristics that Make Us Smart.* Oxford University Press, 1999, pp. 3 – 34.

Gilbert, Nigel, and Klaus Troitzsch. 1999. *Simulation for the Social Scientist.* Philadelphia: Open University Press.

Gilovich, Thomas, Robert Vallone, and Amos Tversky. "The Hot Hand in Basketball: On the Misperception of Random Sequences." *Cognitive Psychology.* 17 (1985).

Gintis, H. Strong reciprocity and human sociality. *J. Theor. Biol.* 206, 169-179 (2000).

Gintis, Herbert. 2000. *Game Theory Evolving.* Princeton: Princeton University Press.

Glaser, William A. 1959. "The Family and Voting Turnout." *Public Opinion Quarterly,* Vol. 23, No. 4. (Winter), 563-570.

Glick, T. F. 1970. *Irrigation and Society in Medieval Valencia.* Cambridge: Harvard University Press.

Green, Donald P., and Ian Shapiro. *Pathologies of Rational Choice Theory.* New Haven: Yale University Press, 1994.

Greenberg, Joseph and Kenneth Shepsle. 1987. "The Effect of Electoral Rewards in Multiparty Competition with Entry." *American Political Science Review,* 81(2):525-538.

Grofman, Bernard, Christian Collet, and Robert Griffin. 1998. "Analyzing The Turnout-Competition Link With Aggregate Cross-Sectional Data." *Public Choice* 95:3-4, 233-246.

Guth, Werner, and Kerstin Pull. 2004. "Will Equity Evolve? An Indirect Evolutionary Approach." *European Journal of Political Economy,* 20:273-282.

Hamilton, W. D. 1964. Genetical evolution of social behavior I and II. *J. Theor. Biol.* 7, 1-52.

Hamilton, W. D. 1975. "Innate social aptitudes of man: an approach from evolutionary genetics." In R. Fox (ed.), *Biosocial Anthropology.* New York: Wiley.

Hansen, Steven, Thomas R. Palfrey, and Howard Rosenthal 1987. "The Relationship Between Constituency Size and Turnout: Using Game Theory to Estimate the Cost of Voting." *Public Choice,* 52:15-33.

Hansson, Ingemar, and Charles Stuart. 1984. "Voting Competitions with Interested Politicians: Platforms do not converge to the preferences of the median." *Public Choice* 44, 431-441.

Hardin, G. The tragedy of the commons. *Science* 162, 1243-1248 (1968).

Henrich, Joseph, and Robert Boyd. 2001, "Why People Punish Defectors," *Journal of Theoretical Biology*, 208:79-89.

Henrich, Joseph, Robert Boyd, Samuel Bowles, Colin Camerer, Ernst Fehr, Herbert Gintis, and Richard McElreath. 2001. "Cooperation, Reciprocity and Punishment in Fifteen Small-scale Societies." *American Economic Review*, 91: 73–78.

Herzberg, Roberta Q. and Rick K. Wilson. 1988. "Results on Sophisticated Voting in an Experimental Setting." *Journal of Politics*, 50:471-86.

Highton Benjamin. 2000. "Residential Mobility, Community Mobility, and Electoral Participation." *Political Behavior* 22:2, 109-120.

Highton, Benjamin. 1997. "Easy Registration and Voter Turnout." *Journal of Politics* 59:2, 565-575.

Hobbes, Thomas. 1651 (1947). *Leviathan.* New York: E. P. Dutton and Co.

Holt, Charles A. and Lisa R. Anderson. 1999. "Agendas and Strategic Voting." *Southern Economic Journal*, January.

Hotelling, Harold. 1929. "Stability in Competition," *The Economic Journal.* 39, 41-57.

Huang, Chi and Todd G Shields. 2000. "Interpretation of Interaction Effects in Logit and Probit Analyses - Reconsidering The Relationship Between Registration Laws, Education, and Voter Turnout." *American Politics Quarterly* 28:1, 80-95.

Huckfeldt, Robert and John Sprague. 1995. *Citizens, Parties, and Social Communication.* New York: Cambridge University Press.

Humphrey, N.K. "The Social Function of Intellect" in P.P.G. Bateson and R.A. Hinde. *Growing Points in Ethology*. Cambridge: Cambridge University Press, 1976, chapter 9, pp. 303 -317.

Isaac, R. M., J. Walker, and S. Thomas. 1984. "Divergent Evidence on Free Riding: An Experimental Examination of Possible Explanations." *Public Choice* 43:113-149.

Iversen, Torben. 1994. "The Logics of Electoral Politics: Spatial, Directional and Mobilizational Effects," *Comparative Political Studies* 27, 155-89.

Jackson, John E. 1983. "Election Night Reporting and Voter Turnout." *American Journal of Political Science*, Vol. 27, No. 4. (Nov.), 615-635..

Jackson, Robert A. 2000. "Differential Influences On Participation in Midterm Versus Presidential Elections." *Social Science Journal* 37:3, 385-402.

Jones, Bradford, Benjamin Radcliff, Charles Taber, and Richard Timpone. 1995. "Condorcet Winners and the Paradox of Voting: Probability Calculation for Weak Preference Orders." *American Political Science Review*, 89:137-144.

Junker, L. L. *Raiding, Trading, and Feasting: The Political Economy of Philippine Chiefdoms* (University of Hawaii Press, Honolulu, 1999).

Kaempfer William H. and Anton D. Lowenberg. 1993. "A Threshold-Model of Electoral Policy and Voter Turnout." *Rationality and Society* 5:1, 107-126.

Kahneman, D., Slovic, P., and Tversky, A. (Eds.). *Judgment Under Uncertainty: Heuristics and Biases.* New York: Cambridge University Press, 1982.

Kahneman, Daniel, and Amos Tversky. "Intuitive Prediction: Biases and Corrective Procedures." *TIMS Studies in Management Science.* 12 (1979), 313 – 327.

Kahneman, Daniel, and Amos Tversky. "On the Psychology of Prediction." *Psychological Review.* 80 (1973), pp. 237 – 251.

Kahneman, Daniel, and Amos Tversky. "On the Reality of Cognitive Illusions." *Psychological Review.* 103 (1996), pp. 582 – 591.

Kahneman, Daniel, and Amos Tversky. "Subjective Probability: A Judgment of Representativeness." *Cognitive Psychology.* 3 (1972), pp. 430 – 454.

Karp, Jeffrey A. and Susan A. Banducci. 2000. "Going Postal: How All-Mail Elections Influence Turnout." *Political Behavior* 22:3, 223-239.

Katz, Daniel, et al. 1954. *Public Opinion and Propaganda.* New York: Henry Holt and Company.

Keeley, L. H. *War Before Civilization; The Myth of the Peaceful Savage* (Oxford University Press, New York, 1997).

Kelly, Jerry S. *Social Choice Theory.* Berlin: Springer-Verlag, 1988.

Kenny, Christopher B. 1992. "Political Participation and Effects from the Social Environment." *American Journal of Political Science* 36:259-267.

Kenrick, D., Sadalla, E., and Keefe, R. "Evolutionary Cognitive Psychology: The Missing Heart of Modern Cognitive Science." In C. Crawford and D. Krebs (Eds.), *Handbook of Evolutionary Psychology*. Mahwah: Lawrence Erlbaum, 1998, pp. 485 – 514.

Key, V.O. 1949. Southern Politics in State and Nation. New York: Knopf.

Kibreab, G. *State Intervention and the Environment in Sudan, 1889-1989* (Edwin Mellen Press, Lewiston, 2002), p. 403.

King, Gary, Michael Tomz, and Jason Wittenberg. 2000. "Making the Most of Statistical Analyses: Improving Interpretation and Presentation." *American Journal of Political Science*, Vol. 44, No. 2 (March): 341-355.

Kingdon, J. 1966. *Candidates for Office: Beliefs and Strategies.* New York: Random House.

Kirchgassner, Gebhard and Anne Meyer zu Himmern. 1997. "Expected Closeness and Turnout: An Empirical Analysis For The German General Elections, 1983-1994." *Public Choice* 91:1, 3-25.

Knack, Stephen. 1992. "Civic norms, social sanctions, and voter turnout." *Rationality and Society* 4:133-156.

Knack, Stephen. 1997. "The Reappearing American Voter: Why Did Turnout Rise in '92?." *Electoral Studies* 16:1, 17-32.

Knack, Stephen. 2001. "Election-Day Registration - The Second Wave." *American Politics Research* 29:1, 65-78.

Kollman, Ken, John H. Miller, and Scott E. Page. 1992. "Adaptive Parties in Spatial Elections," *American Political Science Review* 86:929-937.

Krehbiel, Keith, and Douglas Rivers. 1990. "Sophisticated Voting in Congress: A Reconsideration." *The Journal of Politics*, 52:548-578.

Kreps, David, and Robert Wilson. 1982. "Reputation and Imperfect Information." *Journal of Economic Theory* 27:253-279.

Kreps, David, Paul Milgrom, John Roberts, and Robert Wilson. 1982. "Rational Cooperation in the Finitely Repeated Prisoners' Dilemma." *Journal of Economic Theory* 27:245-252.

Kunce Mitch. 2001. "Pre-Election Polling and The Rational Voter: Evidence From State Panel Data (1986-1998)." *Public Choice* 107:1-2, 21-34.

Kurzban, R., K. McCabe, K. Smith, and B. Wilson. 2001. "Incremental Commitment in a Real-time Public Goods Game." *Personality and Social Psychology Bulletin* 27:1662-1673.

Lazarsfeld, Paul F., Bernard Berelson, and Hazel Gaudet. 1948. *The People's Choice.* New York: Columbia University.

Leal, Donald R. 1996. "Community-Run Fisheries: Avoiding the Tragedy of the Commons." *PERC Policy Series* 7. Bozeman, MT: PERC.

LeBlanc, S. & Register, K. E. *Constant Battles: The Myth of the Peaceful, Noble Savage* (St Martin's Press, London, 2003).

Ledyard, J. 1995. In *The Handbook of Experimental Economics* (eds Kagel, J. H. & Roth A. E.) 111-194. Princeton Univ. Press, Princeton, 1995.

Levi, Margaret. 1988. *Of Rule and Revenue.* University of California Press.

Lindbeck, Assar, and Jorgen W. Weibull 1993. "A Model of Political Equilibrium in a Representative Democracy," *Journal of Public Economics* 51, 195-209.

Lubell, Mark, and John T. Scholz. 2001. "Cooperation, Reciprocity, and the Collective-Action Heuristic." *American Journal of Political Science* 45:160-178.

Lupia, Arthur, and Mathew D. McCubbins. 1998. *The Democratic Dilemma: Can Citizens Learn What They Need to Know?* Cambridge University Press.

Maas, Arthur, and Raymond L. Anderson. 1978. *...and the Desert Shall Rejoice: Conflict, Growth, and Justice in Arid Environment.* Cambridge: The MIT Press.

Mabry, J. B. (ed) *Canals and Communities* (Univ. of Arizona Press, Tucson, 1996).

Mackie, Tom and Richard Rose. 1997. *A Decade of Elections Results: Updating the International Almanac.* Centre for the Study of Public Policy, University of Strathclyde. Glasgow.

Mailath, George J. 1992. "Introduction: Symposium on Evolutionary Game Theory." *Journal of Economic Theory*, 57:259-277.

Mailath, George J. 1998. "Do People Play Nash Equilibrium? Lessons from Evolutionary Game Theory." *Journal of Economic Literature*, 36:1347-1374.

March, James G. "Bounded Rationality, Ambiguity, and the Engineering of Choice." *Bell Journal of Economics*. 9 (Autumn 1978), pp. 587 – 608.

Martin, Lisa L. "The Contributions of Rational Choice: A Defense of Pluralism." *International Security*. 24 (Fall 1999), pp. 74 – 83.

Matsusaka John G. 1993. "Election Closeness and Voter Turnout - Evidence From California Ballot Propositions." *Public Choice* 76:4, 313-334.

May, Robert M. and Brian Martin. 1975. "Voting Models Incorporating Interactions Between Voters." *Public Choice* 22:37-54.

McCloskey, Herbert, Paul J. Hoffmann, and Rosemary O'Hara. 1960. "Issue Conflict and Consensus Among Party Leaders and Followers." *American Political Science Review* 44: 406-427.

McKelvey, Richard. 1976. "Intransitivities in Multidimensional Voting Models and Some Implications for Agenda Control." *Journal of Economic Theory* 12: 472-82.

McKenzie, Richard B., and Gordon Tullock. *The New World of Economics*. Homewood: Richard D. Irwin, 1975.

Meltzer, E. 1994. "Global overview of straddling and highly migratory fish stocks: The nonsustainable nature of high seas fisheries." *Ocean Development and International Law* 25:255-344.

Mitchell, William C. "*The Calculus of Consent*: Enduring Contributions to Public Choice and Political Science." *Public Choice*. 60 (1989), pp. 201 – 210.

Mitchell, William C., and Michael C. Munger. "Economic Models of Interest Groups: An Introductory Survey." *American Journal of Political Science*. 35 (May 1991), pp. 512 – 546.

Mouw, Calvin J., and Michael B. Mackuen. 1992. "The Strategic Agenda in Legislative Politics." *American Political Science Review*, 86:1992.

Mueller, Dennis C. 1989. *Public Choice II*. Cambridge: Cambridge University Press.

Myerson, Roger. 1998. “Population Uncertainty and Poisson Games.” *International Journal of Game Theory* 27:375-392.

Nagler, Jonathan. 1991. "The Effect of Registration Laws and Education On United-States Voter Turnout." *American Political Science Review* 85:4, 1393-1405.

Nalebuff, Barry and Ron Shachar. 1999. “Follow the Leader: A Pivotal Voter Theory”, *American Economic Review*, 89 (3): 525-547.

Nelson, Richard R. 1995. “Recent Evolutionary Theorizing About Economic Change.” *Journal of Economic Literature*, 33:48-90.

Nelson, Richard R., and Sidney G. Winter. 2002. “Evolutionary Theorizing in Economics.” *Journal of Economic Perspectives*, 16:23–46.

Niou, Emerson M.S., and Ordeshook, Peter C. “Return of the Luddites.” *International Security*. 24 (Fall 1999), pp. 84 – 96.

Nossal, Kim R. 1989. “International Sanctions as International Punishment.” *International Organization* 43:301-322.

Oliver, J. Eric. 1996. "The Effects of Eligibility Restrictions and Party Activity On Absentee Voting and Overall Turnout." *American Journal of Political Science* 40:2, 498-513.

Oliver, J. Eric. 1996. "The Effects of Eligibility Restrictions and Party Activity On Absentee Voting and Overall Turnout." *American Journal of Political Science* 40:2, 498-513.

Olsen, Marvin E. 1972. "Social Participation and Voting Turnout: A Multivariate Analysis." *American Sociological Review*, Vol. 37, No. 3. (Jun.), 317-333.

Olson, Mancur. 1965. *The Logic of Collective Action.* Cambridge: Harvard University Press.

Ophuls, W. 1973. “Leviathan or Oblivion.” In H. E. Daly (ed.) ,*Toward a Steady State Economy*, 215-230. San Francisco: Freeman.

Orbell, J. M., van de Kragt, A. & Dawes, R. “Explaining discussion-induced cooperation.” *J. Pers. Soc. Psychol.* 54, 811-819 (1988).

Orbell, John, and Robyn M. Dawes. “A ‘Cognitive Miser’ Theory of Cooperators’ Advantage.” *American Political Science Review.* 85 (June 1991), pp. 515 – 528.

Orbell, John, Tomonori Morikawa, Jason Hartwig, James Hanley, and Nicholas Allen. 2004. “Machiavellian Intelligence as a Basis for the Evolution of Cooperative Dispositions.” *American Political Science Review*, 98:1-16.

Ordeshook, Peter C. *A Political Theory Primer*. New York: Routledge, 1992.

Ordeshook, Peter C. *Game Theory and Political Theory*. Cambridge: Cambridge University Press, 1986.

Ordeshook, Peter C., and Thomas Schwartz. 1987. “Agendas and the Control of Political Outcomes.” *American Political Science Review*, 81:179-200.

Ostrom Elinor, Roy Gardner, and James Walker. 1994. *Rules, Games, and Common-Pool Resources*. Ann Arbor: The University of Michigan Press.

Ostrom, Elinor, and Roy Gardner. 1993. Coping with Asymmetries in the Commons: Self-Governing Irrigation Systems Can Work. *Journal of Economic Perspectives* 7: 93-112.

Ostrom, Elinor. 1990. *Governing the Commons: The Evolution of Institutions for Collective Action*. Cambridge: Cambridge University Press.

Ostrom, Elinor. 1992. “The Rudiments of a Theory of the Origins, Survival, and Performance of Common-Property Institutions.” In D. W. Bromley (ed.) *Making the Commons Work*. San Francisco: ICS Press.

Ostrom, Elinor. 1998. “A Behavioral Approach to the Rational Choice Theory of Collective Action: Presidential Address, American Political Science Association, 1997.” *The American Political Science Review*, 92:1, 1-22.

Palfrey, Thomas R. 1984. “Spatial Equilibrium with Entry.” *Review of Economic Studies* 51: 139-156.

Palfrey, Thomas R., and Howard Rosenthal. 1985. “Voter Participation and Strategic Uncertainty.” *American Political Science Review* 79:62-78.

Patton, J. Q. in *Adaptation and Human Behaviour: An Anthropological Perspective* (eds. Cronk, L., Chagnon, N. & Irons, W.) 417-436 (Aldine de Gruyter, Hawthorne, New York, 2000).

Peltzman, Samuel. 1984. "Constituent Interest and Congressional Voting." *Journal of Law and Economics* 27:181-210.

Peltzman, Samuel. 1984. "Constituent Interest and Congressional Voting." *Journal of Law and Economics* 27:181-210.

Plott, Charles R. 1983. "Externalities and Corrective Policies in Experimental Markets." *Economic Journal* 93:106-127.

Plott, Charles R. and Michael E. Levine. 1978. "A Model of Agenda Influence on Committee Decisions." *American Economic Review* 68:146-60.

Plott, Charles. 1967. "A Notion of Equilibrium Under Majority Rule," *American Economic Review*. 57: 787-806.

Plott, Charles. 1967. "A Notion of Equilibrium Under Majority Rule," *American Economic Review*. 57: 787-806.

Poole, Keith T., and Howard Rosenthal 1984. "The Polarization of American Politics." *The Journal of Politics* 46(4), 1061-1079.

Poulton, E. C. *Behavioral Decision Theory: A New Approach*. Cambridge: Cambridge University Press, 1994.

Price, M. E., Cosmides, L. & Tooby, J. Punitive sentiment as an anti-free rider psychological device. *Evol. Hum. Behav.* 23, 203-231 (2002).

Quattrone, George A., and Amos Tversky. "Contrasting Rational and Psychological Analysis of Political Choice." *American Political Science Review.* 82 (September 1988), pp.719 – 736.

Quattrone, George A., and Amos Tversky. 1988. "Contrasting Rational and Psychological Analysis of Political Choice." *American Political Science Review.* 82, 719– 736.

Rhine, Staci L. 1995. "Registration Reform and Turnout Change in The American States." *American Politics Quarterly* 23:4, 409-426.

Riker, William H. 1980. "Implications from the Disequilibrium of Majority Rule." *American Political Science Review*, 74:432-46.

Riker, William H. *Liberalism Against Populism*. Illinois: Waveland Press, 1982.

Riker, William H. *The Theory of Political Coalitions*. New Haven: Yale University Press, 1963.

Riker, William H., and Peter C. Ordeshook. 1968. "A Theory of the Calculus of Voting." *American Political Science Review* 62:25-42.

Riker. William H. 1986. *The Art of Political Manipulation.* New Haven: Yale University Press.

Roemer, John E. 2001. *Political Competition: Theory and Applications.* Cambridge: Harvard University Press.

Rosenstone, Steven J. and Raymond E. Wolfinger. 1978. "The Effect of Registration Laws On Voter Turnout." *American Political Science Review*, Vol. 72, No. 1. (Mar.), 22-45..

Ruttan, L. 1998. "Closing the Commons: Cooperation for Gain or Restraint?" *Human Ecology* 26:43-66.

Sakurai, T., Y. Kijima, R. Pokharel, S. Rayamajhi, and K. Otsuka. "Timber Forest Management in Nepal and Japan." In K. Otsuka and F. Place (eds.), *Land Tenure and Natural Resource Management: A Comparative Study of Agrarian Communities in Asia and Africa.* Baltimore: Johns Hopkins University Press.

Samuelson, Larry. 1997. *Evolutionary Games and Equilibrium Selection.* Cambridge: The MIT Press.

Schelling, Thomas C. *The Strategy of Conflict.* Cambridge, 1960 (1980 ed.).

Sethi, R., and Somanathan E. 1996. "The Evolution of Social Norms in Common Property Resource Use." *American Economic Review* 86:766-788.

Shavell, Steven. 1987. "The Optimal Use of Nonmonetary Sanctions as a Deterrent". *American Economic Review* 77:584-592.

Shepsle, Kenneth A. "Institutional Arrangements and Equilibrium in Multidimensional Voting Models." *American Journal of Political Science.* 23, pp. 27 – 59.

Shepsle, Kenneth A., and Barry R. Weingast. "Positive Theories of Congressional Institutions." *Legislative Studies Quarterly.* 19 (May 1994), pp. 149 – 179.

Shepsle, Kenneth A., and Barry R. Weingast. 1984. "Uncovered Sets and Sophisticated Voting Outcomes with Implications for Agenda Institutions." *American Journal of Political Science*, 28:49-74.

Shepsle, Kenneth A., and Mark S. Bonchek. *Analyzing Politics: Rationality, Behavior, and Institutions.* New York: W.W.Norton, 1997.

Simon, Herbert. "Human Nature in Politics: The Dialogue of Psychology with Political Science." *American Political Science Review.* 79 (June 1985), pp. 293 – 304.

Simon, Herbert. 1982. *Models of Bounded Rationality.* Cambridge: MIT Press.

Simon, Herbert. 1985. "Human Nature in Politics: The Dialogue of Psychology with Political Science." *American Political Science Review.* 79, 293 – 304.

Simon, Herbert. 1990. A mechanism for social selection and successful altruism. *Science* 250, 1665-1668.

Simon, Herbert. *Models of Bounded Rationality.* Cambridge: MIT Press, 1982.

Slovic, Paul, G. Fischhoff, and S. Lichtenstein. "Facts Versus Fears: Understanding Perceived Risk." In Kahneman, D., Slovic, P., and Tversky, A. (Eds.). *Judgment Under Uncertainty: Heuristics and Biases.* New York: Cambridge University Press, 1982, pp. 463 – 489.

Slovic, Paul. *Smoking: Risk, Perception, and Policy.* Sage Publications, 2001.

Smith, Maynard. 1982. *Evolution and the Theory of Games*. Cambridge: Cambridge University Press.

Smith, R. J. 1981. Resolving the Tragedy of the Commons by Creating Private Property Rights in Wildlife. *CATO Journal* 1:439-468.

Sober, E. & Wilson, D. S. *Unto Others: The Evolution and Psychology of Unselfish Behaviour* (Harvard Univ. Press, Cambridge, Massachusetts, 1998).

Southwell, Priscilla L. and Justin Burchett. 2000b. "The Effect of All-Mail Elections On Voter Turnout." *American Politics Quarterly* 28:1, 72-79.

Squire, Peverill, Raymond E. Wolfinger, and David P. Glass. 1987. "Residential Mobility and Voter Turnout (in Articles)." *American Political Science Review* 81:1. (Mar.), 45-66.

Stahl, Dale O. 1993. "Evolution of Smart Players." *Games and Economic Behavior,* 5:604-617.

Stevens, Joe. B. *The Economics of Collective Choice*. Boulder: Westview Press, 1993.

Stevenson, Glenn G. 1991. *Common Property Economics: A General Theory and Land Use Applications*. Cambridge: Cambridge University Press.

Stolper, Wolfgang and Paul A. Samuelson. 1941. "Protection and Real Wages." *Review of Economic Studies*, 9:58-73.

Straits, Bruce C. 1990. "The Social-Context of Voter Turnout." *Public Opinion Quarterly* 54:1, 64-73.

Swinkels, Jeroen M. 1993. "Adjustment Dynamics and Rational Play in Games." *Games and Economic Behavior*, 5:455-484.

Tang, S. Y. *Institutions and Collective Action* (Institute for Contemporary Studies, San Francisco, 1992).

Tooby, J. & Cosmides, L. in Institute for Evolutionary Studies, Technical Report #88-1 (University of California at Santa Barbara, Santa Barbara, 1998).

Tooby, John, and Leda Cosmides. 1992. "The Psychological Foundations of Culture" in Jerome H. Barkow, Leda Cosmides, and John Tooby. *The Adapted Mind: Evolutionary Psychology and the Generation of Culture*. New York: Oxford University Press. Pp. 19-136.

Tooby, John, and Leda Cosmides. Evolutionary Psychology Primer, 1997. http://www.psych.ucsb.edu/research/cep/primer.html

Trivers, Robert. 1971. "The evolution of reciprocal altruism." *Quarterly Review of Biology*, 46:35-57.

Tversky, Amos, and Daniel Kahneman. "Availability: A Heuristic for Judging Frequency and Probability." *Cognitive Psychology*. 5 (1973), pp. 207 – 232.

Tversky, Amos, and Daniel Kahneman. "Causal Schemas in Judgments Under Uncertainty" in M. Fishbein (ed.), *Progress in Social Psychology*. Hillsdale: Erlbaum, 1980, pp. 84 – 98.

Tversky, Amos, and Daniel Kahneman. "Extensional Versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgment." *Psychological Review*. 91 (1983), pp. 293 – 315.

Umbeck, John R. 1981. *A Theory of Property Rights: With Application to the California Gold Rush*. Ames: The Iowa State University Press.

Vanberg, V. J. and Congleton, R. D. (1992) "Rationality, Morality and Exit." *American Political Science Review*, 86: 418-431.

Vega-Redondo, Fernando. 1993. "Competition and Culture in an Evolutionary Process o Equilibrium Selection: A Simple Example." *Games and Economic Behavior*, 5:618-631.

Veit, Peter G., Clement Dorm-Adzobu, and Okyeame Ampadu-Agyei. 1991. *Community Institutions in Resource Management: Agroforestry by Mobisquads in Ghana.* World Resources Institute.

Walt, Stephen M. "Rigor or Rigor Mortis? Rational Choice and Security Studies." *International Security.* 23 (Spring 1999), pp. 5 – 48.

Weber, M. *From Max Weber: Essays in Sociology* (Oxford Univ. Press, Oxford, 1958).

Weibull, Jorgen W. 1996. *Evolutionary Game Theory.* Cambridge: The MIT Press.

Weibull, Jorgen W. 1998. "What Have I Learned From Evolutionary Game Theory So Far?" I.U.I. Working Papers, n. 487.

Weingast, Barry. 1979. "A Rational Choice Perspective on Congressional Norms." *American Journal of Political Science. 23*, pp. 245 – 262.

Wilson, M. L. & Wrangham, R. W. Intergroup Relations in Chimpanzees. *Annual Review of Anthropology* 32, 363-392 (2003).

Wilson, Rick K. 1986. "Forward and Backward Agenda Procedures: Committee Experiments on Structurally Induced Equilibrium." *Journal of Politics,* 48:390-409.

Wittman, Donald. 1977. "Candidates with Policy Preferences: A Dynamic Model." *Journal of Economic Theory* 14, 180-189.

Wolfinger, Raymond and Steven Rosenstone. 1980. *Who Votes?* New Haven: Yale University.

Yamagishi, T. 1986. "The provision of a sanctioning system as a public good." *Journal of Personality and Social Psychology* 51:110-116.

Young, H.P. 1988. "Condorcet's Theory of Voting." *American Political Science Review,* 82:1230-44.